IRE
Transactions
on ANTENNAS and PROPAGATION

Volume AP-9 MAY, 1961 Number 3

*Published Bimonthly*

## In This Issue



PUBLISHED BY THE
rofessional Group on Antennas and Propagation

IRE TRANSACTIONS® PGAP IS A PUBLICATION DEVOTED TO EXPERIMENTAL AND THEORETICAL PAPERS ON RADIO ANTENNAS, ON GUIDED OR UNGUIDED PROPAGATION OF RADIO WAVES, AND ON ALLIED FIELDS OF RADIO PHYSICS SUCH AS RADIO ASTRONOMY

***MANUSCRIPTS*** should be submitted to Sidney A. Bowhill, *Editor*, 222 Electrical Engineering, Pennsylvania State University, University Park, Pa. Manuscripts should be original typewritten copy, double-spaced, plus one carbon copy and two sets of copies of illustrations. Original illustrations will be called for if the paper is accepted. References should appear as footnotes and include author's name, title, journal, volume, initial and final page numbers, and date.

***CONTRIBUTIONS,*** which should average 15 double-spaced typewritten pages in length, are subjected to review by the *Associate Editors* and their readers. Each paper must have a summary of less than 200 words.

***COMMUNICATIONS*** should not exceed five double-spaced typewritten pages in length, together with not more than three illustrations. Accepted at the Editor's discretion, they appear in the first available issue.

***NEWS ITEMS*** concerning PGAP members and group activities should be sent to the *News Editor*, R. C. Hansen, Space Technology Laboratories, P.O. Box 95001, Los Angeles 45, Calif.

***ORIGINAL ILLUSTRATIONS*** should be submitted as follows: All line drawings (graphs, charts, block diagrams, cutaways, etc.) should be inked uniformly and ready for reproduction. If commercially printed grids are used in graph drawings, author should be sure printer's ink is of a color that will reproduce. Photographs should be glossy prints. Call-outs or labels should be marked on a registered tissue overlay, not on the illustration itself. No illustration should be larger than 8 x 10 inches. Lettering on illustrations must have height at least two per cent of the illustration width.

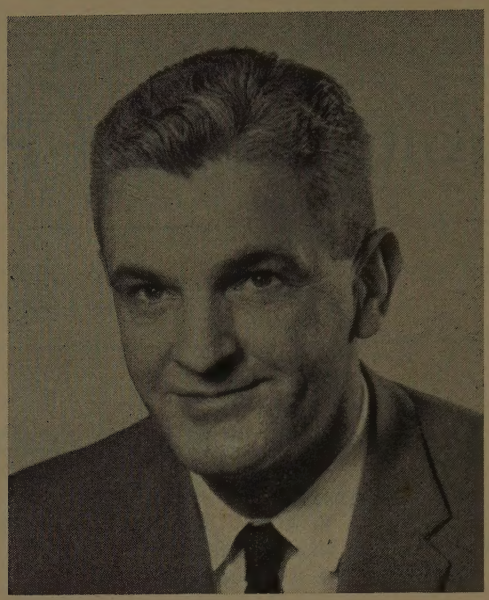
***Copies can be purchased from*** **the INSTITUTE OF RADIO ENGINEERS, 1 East 79 St., New York 21, N.Y.** Individual copies of this issue, and all available back issues, except Vols. AP-5, No. 1; AP-6, No. 1; AP-7, Special Supplement, may be purchased at the following prices: IRE members (one copy) $2.25, libraries and colleges $3.25, all others $4.50. Yearly subscription rate: non-members $17.00; colleges and public libraries $10.00. IRE TRANSACTIONS ON ANTENNAS AND PROPAGATION.  Printed in U.S.A. Printed by George Banta Co., Inc., Curtis Reed Plaza, Menasha, Wisconsin.

Second-class postage paid at MENASHA, WISCONSIN, and additional mailing offices under the act of August 24, 1912. Acceptance for mailing at a special rate of postage is provided for in the act of February 28, 1925, embodied in Paragraph 4, Section 412, P. L. & R., authorized October 26, 1927.

# IRE Transactions on Antennas and Propagation

Volume AP-9 | MAY, 1961 | Number 3
*Published Bimonthly*

## TABLE OF CONTENTS



# John T. Bolljahn, 1918-1960

John T. Bolljahn was born in Oakland, Calif., on June 10, 1918, and passed away at Palo Alto, Calif., on June 25, 1960, in his forty-second year. He was the son of Herman Bolljahn, a construction engineer who had emigrated to California from Germany as a young man, and Lillian A. Smith, who was born in Glendale, Calif. He married Harriette M. Pedrotti on September 19, 1942, and fathered three children, a son, Frederick, who passed away in 1948 at the age of 2 years, and daughters Mary and Claire, born in 1947 and 1951. He received his education at the University of California in Berkeley, from which he received the B.S. degree in May, 1941, and the Ph.D. degree in January, 1950, both degrees in electrical engineering.

Dr. Bolljahn was on the staff of the Naval Research Laboratory, Washington, D. C., from August, 1941, through December, 1945, as Antenna Group Leader of the Aircraft Section, Radio Division. He returned to the University of California in January, 1946, to continue his graduate studies and to participate, as a staff engineer, in the antenna research program conducted at the University under the auspices of the Office of Naval Research. During this interval he served the University also as a lecturer in electrical engineering. In October, 1949, he joined the newly-formed Aircraft Radiation Systems Laboratory of the Stanford Research Institute in Menlo Park, Calif., as assistant supervisor. He assumed a series of increasing responsibilities at the Institute, and was named Assistant Director of Engineering Research in 1956. He retained an association with the Institute until his death. In 1956, with three others, he formed Granger Associates, serving as a director of that corporation and, in March, 1960, became its Executive Vice President.

Dr. Bolljahn was a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He was named a Fellow of the IRE in March, 1960, "for his contributions to the theory and practical design of antennas." He served at various times on scientific advisory committees to several government bodies and industrial groups. He was active in the affairs of the IRE in many capacities, and served on the Administrative Committee of the Professional Group on Antennas and Propagation from 1953 to 1957.

Jack Bolljahn's efforts in behalf of his profession were numerous, of exceptional quality, and of great significance to the advancement of the radio art. He made outstanding contributions to the theory of electrically small antennas and to practical methods for their design and test. He added important findings to the theory of electromagnetic scattering. His work on the properties of coupled, distributed elements led to fundamental advances in the design of filters and transmission components. His understanding of the special problems of antennas for aircraft brought to every worker in that field a unifying clarification of the basic phenomena involved and a wealth of new, intensely practical, design approaches. He wrote 29 technical reports on a wide range of topics. Ten papers reporting his research were published in scientific journals, and he made many presentations to technical meetings. In addition, he was in the process of completing a number of additional papers. One of these "Synthesis of Modulated Corrugated Surface-Wave Structures," appears in this issue of the TRANSACTIONS. It is hoped that others can be published at a later date.

Any outline of Dr. Bolljahn's career and his contributions to the literature cannot begin to convey the breadth and depth of the man. His personal and professional influence extended beyond his immediate associates to hundreds of others to whom his wisdom, his analytical skills, his keen perception, his selfless dedication, his warm and generous spirit, brought new appreciations of the challenge and the rewards of his profession. It is a privilege to have known him and an honor to have been his friend. —J. V. N. GRANGER

# The John T. Bolljahn Memorial Award

By action of the Administrative Committee of the Professional Group on Antennas and Propagation, the annual prize for the best paper published in the Transactions of the Professional Group on Antennas and Propagation is to be known as the John T. Bolljahn Memorial Award. The purpose of the award, in memory of the late Dr. John T. Bolljahn, is to recognize outstanding papers in the field of antennas and propagation by the award of a cash prize, and thereby to stimulate increased interest in the field.

In order to make the amount of the award substantial enough to serve as a stimulus for the generation of meritorious papers, the $200 presently allotted is to be increased by the proceeds from a fund supported by contributions. In this way it is hoped that the honorarium could total $500 annually. A fund-raising committee has been appointed by Dr. E. C. Jordan, Chairman of the Administrative Committee. This Committee, acting on its own behalf and not as agents of the Institute of Radio Engineers, is in the process of soliciting contributions to this fund, which then will be given as a gift to the IRE for the purpose of the John T. Bolljahn Memorial Award. Co-chairmen of the fund-raising committee are as follows:

M. Katzin
5001 College Avenue
College Park, Md.

J. V. N. Granger
974 Commercial Street
Palo Alto, Calif.

# contributions

## Synthesis of Modulated Corrugated Surface-Wave Structures*

J. T. BOLLJAHN†, FELLOW, IRE

*Summary*—**An exact procedure for the design of a surface that will support a prescribed group of surface waves simultaneously is described. The surfaces resulting from the application of this procedure are found to be modulated (*i.e.*, to contain periodic waves in the direction of propagation) and to have surface impedance values that vary periodically in the direction of propagation. Finite-length surfaces of this type are hence reminiscent of Simon's "cigar antenna." Several surface designs are presented, and some general properties of the surfaces and the composite waves they support are discussed. Certain limitations on the design procedure are described.**

### I. Introduction

THIS PAPER describes a method for designing a modulated corrugated surface-wave structure to support a given set of surface-wave modes. Interest in this problem stems from recent work by Simon and Biggi,[1] Thomas and Zucker,[2] and Pease.[3]

Simon and Biggi's work[1] describes the so-called cigar antenna, which is a disk-on-rod structure in which the disk diameters and/or spacings are varied periodically in the axial direction. Thomas and Zucker[2] consider the synthesis of desired radiation patterns using such antenna configurations, on the theory that the aperture distribution along the antenna may be expressed as the superposition of a number of discrete modes, each in the form of a simple trapped wave. Pease[3] considers the approximate solutions for waves bound to a dielectric slab in which the dielectric constant is modulated periodically in the direction of propagation. The latter two papers[2,3] consider not only end-fire radiators, but also apertures which radiate in other than the end-fire direction.

The approach used here is exact in the sense that it determines the characteristics (*i.e.*, shape and surface impedance) of that surface (or family of surfaces) which satisfies the boundary conditions in the presence of any prescribed set of trapped waves. The analysis, like those in Thomas and Zucker, and in Pease, is applicable to a two-dimensional model. A more interesting solution from the practical design standpoint would be that for a cylindrical end-fire structure. Unfortunately, the technique developed is not valid for the cylindrical geometry. This limitation is discussed in the Appendix.

* Received by the PGAP, July 13, 1960. The work described in this paper was sponsored by the AF Cambridge Res. Ctr. under Contract AF 19(604)-3502 with Stanford Res. Inst., Menlo Park, Calif.

† Stanford Res. Inst., Menlo Park, Calif. Granger Associates, Palo Alto, Calif., at the time of his death, June 26, 1960.

[1] J. C. Simon and V. Biggi, "Un nouveau type d'aèrien et son application à la transmission de télévision à grand distance," *L'Onde Elec.*, vol. 332, pp. 1–16; November, 1954. (In French.)

[2] A. S. Thomas and F. J. Zucker, "Radiation from modulated surface-wave structures—I," 1957 IRE National Convention Record, pt. 1, pp. 153–160.

[3] R. L. Pease, "Radiation from modulated surface-wave structures—II," 1957 IRE National Convention Record, pt. 1, pp. 161–165.

A further limitation is that the analysis is inherently applicable only to the corrugated-surface type of structure and not to the dielectric slab form of trapped-wave surface. This limitation arises since the analysis postulates that there is no energy flow across the interface between the guiding structure and the free-space region above this structure. In a dielectric slab radiator supporting more than one mode, energy can travel in the direction of propagation both within the slab and in the medium above the slab, and a periodic interchange of energy can occur between the two regions across the interface. The restriction that the interface be a surface across which no energy passes is therefore not applicable in this case. In a corrugated-surface structure, there is no energy flow below the interface, and hence the restriction is a valid one.

## II. Analysis

We shall confine attention to the two-dimensional case (Fig. 1) for which we may represent the simple trapped

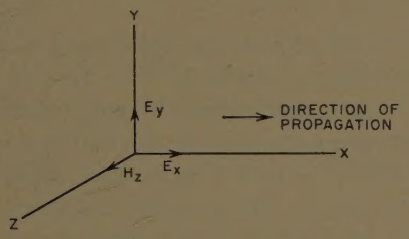



Fig. 1—Coordinate system.

wave by the field components,

$$E_x = \frac{jB}{\omega\epsilon} e^{-\alpha y - j\beta x}$$

$$E_y = \frac{B}{\omega\epsilon}\frac{\beta}{\alpha} e^{-\alpha y - j\beta x}$$

$$H_z = \frac{B}{\alpha} e^{-\alpha y - j\beta x}$$

$$\alpha^2 = \beta^2 - k^2. \tag{1}$$

We wish now to postulate a group of such waves having different $\alpha$'s and $\beta$'s, but the same frequency, and to find the flow lines represented by the directions of the real part of the complex Poynting vector. These flow lines or surfaces are the surfaces across which there is no energy flow, and are hence the surfaces which, if designed to present the proper surface reactance, can support the set of waves postulated. The imaginary part of the complex Poynting vector can be used, as described later, to define the surface reactance required along the surfaces defined above.

For a group of $n$ waves, we may write

$$\begin{aligned} \boldsymbol{E}\times\boldsymbol{H}^* &= \bar{a}_x E_y H_z^* - \bar{a}_y E_x H_z^* \\ &= \bar{a}_x \frac{1}{\omega\epsilon}\sum_n B_n \frac{\beta_n}{\alpha_n} e^{-\alpha_n y - j\beta_n x} \cdot \sum_n \frac{B_n}{\alpha_n} e^{-\alpha_n y + j\beta_n x} \\ &\quad - \bar{a}_y \frac{j}{\omega\epsilon}\sum_n B_n e^{-\alpha_n y - j\beta_n x} \cdot \sum_n \frac{B_n}{\alpha_n} e^{-\alpha_n y + j\beta_n x} \end{aligned} \tag{2}$$

Separating real and imaginary parts,[4]

$$\begin{aligned} &\omega\epsilon \operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*) \\ &= \bar{a}_x\left[\sum_{p=1}^{n}\left(\frac{B_p}{\alpha_p}\right)^2 \beta_p e^{-2\alpha_p y} \right. \\ &\quad \left. + \sum_{p=1}^{n}\sum_{q=1}^{p-1}\frac{B_p B_q}{\alpha_p\alpha_q}(\beta_q+\beta_q)e^{-(\alpha_p+\alpha_q)y}\cos(\beta_p-\beta_q)x\right] \\ &\quad + \bar{a}_y \sum_{p=1}^{n}\sum_{q=1}^{p-1}\frac{B_p B_q}{\alpha_p\alpha_q}(\alpha_q-\alpha_p)e^{-(\alpha_p+\alpha_q)y}\sin(\beta_p-\beta_q)x \end{aligned} \tag{3}$$

$$\begin{aligned} &\omega\epsilon \operatorname{Im}(\boldsymbol{E}\times\boldsymbol{H}^*) \\ &= \bar{a}_x \sum_{p=1}^{n}\sum_{q=1}^{p-1}\frac{B_p B_q}{\alpha_p\alpha_q}(\beta_q-\beta_p)e^{-(\alpha_p+\alpha_q)y}\sin(\beta_p-\beta_q)x \\ &\quad + \bar{a}_y\left[\sum_{p=1}^{n}\frac{B_p}{\alpha_p}e^{-2\alpha_p y} \right. \\ &\quad \left. + \sum_{p=1}^{n}\sum_{q=1}^{p-1}\frac{B_p B_q}{\alpha_p\alpha_q}(\alpha_q+\alpha_p)e^{-(\alpha_p+\alpha_q)y}\cos(\beta_p-\beta_q)x\right]. \end{aligned} \tag{4}$$

Specializing to the case $n=3$, and taking $(\beta_1-\beta_2) = (\beta_2-\beta_3) = \delta$, the expressions become

$$\begin{aligned} &\omega\epsilon \operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*) \\ &= \bar{a}_x\left\{B_1^2\frac{\beta_1}{\alpha_1^2}e^{-2\alpha_1 y} + B_2^2\frac{\beta_2}{\alpha_2^2}e^{-2\alpha_2 y} + B_3^2\frac{\beta_3}{\alpha_3^2}e^{-2\alpha_3 y}\right. \\ &\quad + \cos\delta x\left[B_1B_2\frac{(\beta_1+\beta_2)}{\alpha_1\alpha_2}e^{-(\alpha_1+\alpha_2)y}\right. \\ &\quad \left. + B_2B_3\frac{(\beta_2+\beta_3)}{\alpha_2\alpha_3}e^{-(\alpha_2+\alpha_3)y}\right] \\ &\quad \left. + \cos 2\delta x\left[B_1B_3\frac{(\beta_1+\beta_3)}{\alpha_1\alpha_3}e^{-(\alpha_1+\alpha_3)y}\right]\right\} \\ &\quad + \bar{a}_y\left\{\sin\delta x\left[B_1B_2\frac{(\alpha_2-\alpha_1)}{\alpha_1\alpha_2}e^{-(\alpha_1+\alpha_2)y}\right.\right. \\ &\quad \left. + B_2B_3\frac{(\alpha_3-\alpha_2)}{\alpha_2\alpha_3}e^{-(\alpha_2+\alpha_3)y}\right] \\ &\quad \left. + \sin 2\delta x\left[B_1B_3\frac{(\alpha_3-\alpha_1)}{\alpha_1\alpha_3}e^{-(\alpha_1+\alpha_3)y}\right]\right\} \end{aligned} \tag{5}$$

[4] A more general formulation may be made in which the relative phases of the various modes at $x=0$ are arbitrary. The cases considered in this report, however, are amply covered by the more limited (and compact) expressions given.

$$\omega\epsilon \operatorname{Im}(\boldsymbol{E}\times\boldsymbol{H}^*)$$

$$= -\bar{a}_x\left\{\sin\delta x\left[B_1B_2\frac{(\beta_1-\beta_2)}{\alpha_1\alpha_2}e^{-(\alpha_1+\alpha_2)y}+B_2B_3\frac{(\beta_2-\beta_3)}{\alpha_2\alpha_3}e^{-(\alpha_2+\alpha_3)y}\right]+\sin 2\delta x\left[B_1B_3\frac{(\beta_1-\beta_3)}{\alpha_1\alpha_3}e^{-(\alpha_1+\alpha_3)y}\right]\right\}$$

$$-\bar{a}_y\left\{\frac{B_1^2}{\alpha_1}e^{-2\alpha_1y}+\frac{B_2^2}{\alpha_2}e^{-2\alpha_2y}+\frac{B_3^2}{\alpha_3}e^{-2\alpha_3y}+\cos\delta x\left[B_1B_2\frac{(\alpha_1+\alpha_2)}{\alpha_1\alpha_2}e^{-(\alpha_1+\alpha_2)y}+B_2B_3\frac{(\alpha_2+\alpha_3)}{\alpha_2\alpha_3}e^{-(\alpha_2+\alpha_3)y}\right]+\cos 2\delta x\left[B_1B_3\frac{(\alpha_1+\alpha_3)}{\alpha_1\alpha_3}e^{-(\alpha_1+\alpha_3)y}\right]\right\}. \quad (6)$$

To calculate the surfaces across which no energy flows, we note that since there are no sources of energy in the region under consideration,

$$\bar{\nabla}\cdot\operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*)\equiv 0.$$

Assuming that the surfaces sought are of the form shown in Fig. 2 (there will actually be a family of such surfaces

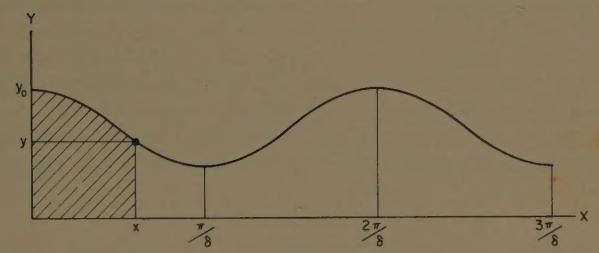



Fig. 2—Application of Gauss's theorem to surface.

corresponding to different values of $y_0$), we may apply Gauss's theorem to the region shown cross-hatched and obtain

$$-\int_0^{y_0}\operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*)_x\Big|_{x=0}dy-\int_0^{x}\operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*)_y\Big|_{y=x_0}dx+\int_0^{y}\operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*)_x\Big|_{x=x_0}dy=0. \quad (7)$$

Substituting from (3) and (4) and carrying out the integrations, we find that

$$B_1^2\frac{\beta_1}{2\alpha_1^3}e^{-2\alpha_1y}+B_2^2\frac{\beta_2}{2\alpha_2^3}e^{-2\alpha_2y}+B_3^2\frac{\beta_3}{2\alpha_3^3}e^{-2\alpha_3y}+\cos\delta x\left[\frac{B_1B_2}{\alpha_1\alpha_2}\left(\frac{\beta_1+\beta_2}{\alpha_1+\alpha_2}\right)e^{-(\alpha_1+\alpha_2)y}+\frac{B_2B_3}{\alpha_2\alpha_3}\left(\frac{\beta_2+\beta_3}{\alpha_2+\alpha_3}\right)e^{-(\alpha_2+\alpha_3)y}\right]+\cos 2\delta x\frac{B_1B_3}{\alpha_1\alpha_3}\left(\frac{\beta_1+\beta_3}{\alpha_1+\alpha_3}\right)e^{-(\alpha_1+\alpha_3)y}$$

$$=B_1^2\frac{\beta_1}{2\alpha_1^3}e^{-2\alpha_1y_0}+B_2^2\frac{\beta_2}{2\alpha_2^3}e^{-2\alpha_2y_0}+B_3^2\frac{\beta_3}{2\alpha_3^3}e^{-2\alpha_3y_0}+\frac{B_1B_2}{\alpha_1\alpha_2}\left(\frac{\beta_1+\beta_2}{\alpha_1+\alpha_2}\right)e^{-(\alpha_1+\alpha_2)y_0}+\frac{B_2B_3}{\alpha_2\alpha_3}\left(\frac{\beta_2+\beta_3}{\alpha_2+\alpha_3}\right)e^{-(\alpha_2+\alpha_3)y_0}+\frac{B_1B_3}{\alpha_1\alpha_3}\left(\frac{\beta_1+\beta_3}{\alpha_1+\alpha_3}\right)e^{-(\alpha_1+\alpha_3)y_0}. \quad (8)$$

For a given value of $y_0$, the right-hand member of this equation becomes a constant. The desired surface for that value of $y_0$ is determined by substituting values of $x$ into the equation and finding the corresponding values of $y$ which lead to satisfaction of the equation.

Up to this point, the surfaces discussed have been defined only as mathematical surfaces with fields both above and below them. If we wish to place a physical boundary on one of these surfaces which will retain just the fields above the surface, it is clear that this boundary must present a surface reactance which varies from point to point in accordance with the ratio of the tangential $E$ field to the tangential $H$ field along the boundary. Thus

$$X_s=\frac{E_t}{H_t}=\frac{\operatorname{Im}(\boldsymbol{E}\times\boldsymbol{H}^*)\times\operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*)}{|H|^2|\operatorname{Re}(\boldsymbol{E}\times\boldsymbol{H}^*)|}. \quad (9)$$

Surface shapes and reactance functions have been calculated for a number of different wave groups, and these are shown in Fig. 3(a)–3(d). It is noted that the first of these is designed for a group of three waves while the others are all for two-wave groups. The basic objective of an end-fire antenna employing such a guiding surface is to permit the launching of a wave having $v/c$ slightly less than unity on a surface having a more tightly trapped-wave system than is obtained on a single-mode surface for the same value of $v/c$. It is noted in these curves that in most cases the average value of surface impedance is considerably higher than that of the surface required to support the fastest wave alone. Since the attenuation of the fields normal to the surface is, to a good approximation, proportional to the local surface

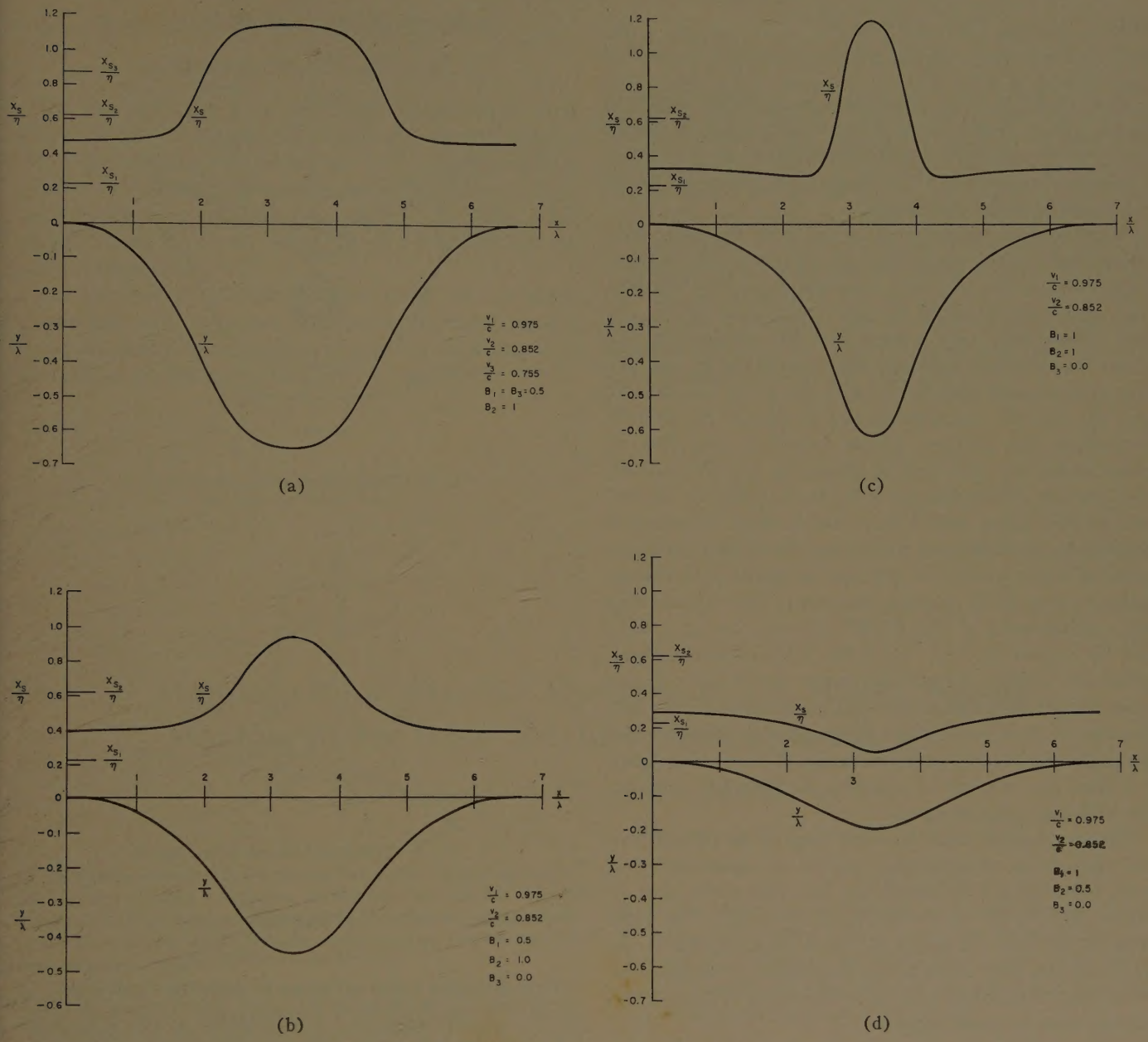



Fig. 3—Modulated corrugated-surface design.

reactance, it follows that the objective of achieving a relatively tightly bound wave system is achievable with this type of structure. The surface reactance reaches a maximum in most cases in the troughs of the modulated surfaces. It is seen, therefore, that in these cases the feed system for such an antenna should be placed in a trough in order to couple most effectively to the wave system to be excited. This is not a general result for all surfaces, however, as is shown by the results in Fig. 3(d).

## III. Some Properties of Surfaces and Fields

A question of interest in connection with the design of such surfaces is: Can a plane-corrugated surface (*i.e.*, one without undulations) be designed to support a finite group of trapped waves by simply varying the surface impedance in the direction of propagation? It is readily shown with the formulation used here that the answer is no. Referring to (3), the $y$ component of Re $(\boldsymbol{E}\times\boldsymbol{H}^*)$ for a group of $n$ waves at a fixed value of $y$ is seen to be of the form

$$\text{Re}\,(\boldsymbol{E}\times\boldsymbol{H}^*)_y\,\big|_{y=\text{const.}} = \frac{1}{\omega\epsilon}\sum_{p=1}^{n}\sum_{q=1}^{p-1}\frac{B_p'B_q'}{\alpha_p\alpha_q}(\alpha_q-\alpha_p)\sin(\beta_p-\beta_q)x, \quad (10)$$

where

$$B_p' = B_p e^{-\alpha_p y},$$

and

$$B_q' = B_q e^{-a_q y}.$$

If $n$ is finite, this expression is a finite series in $\sin(\beta_p - \beta_q)x$. Since there is no combination of coefficients which can reduce this expression to zero for all $x$, it follows that there is no finite group of waves for which the required surface is planar.

Another point of interest in fields of the type considered here is the occurrence of regions in which the $x$ component of Re $(\boldsymbol{E}\times\boldsymbol{H}^*)$ reverses direction. The fact that such regions occur is easily demonstrated in the case where only two waves exist by considering the phase relationships between the total fields $E_y$ and $H_z$, first at $x=0$ and then at $x=\pi/\delta$. At $x=0$, the two $E_y$ and the two $H_z$ components are in phase. At $x=\pi/\delta$, because of the different phase velocities, the two $E_y$ components are 180° out of phase, as are the two $H_z$ components. Since the ratio $E_y/H_z$ is different for the two waves, there will always be a range of relative amplitudes of these two waves for which the resultant value of $E_y$ at $x=\pi/\delta$ is 180° out of phase with the resultant value of $H_z$ at the same point. This means that the $x$ component of Re $(\boldsymbol{E}\times\boldsymbol{H}^*)$ is negative at $x=\pi/\delta$ and for values of $y$ for which the relative wave amplitudes are within the proper range. Since the relative wave amplitudes will go through all possible values as a function of $y$, with the wave having the larger $\alpha$ becoming negligible at large positive values of $y$ and vice versa, it is seen that regions of negative Re $(\boldsymbol{E}\times\boldsymbol{H}^*)_x$ will occur with any pair of waves.

The characteristics of the flow lines of Re $(\boldsymbol{E}\times\boldsymbol{H}^*)$ in the vicinity of this reversal of the $x$ component has been investigated in detail for the two-wave case in which $B_1=B_2=1$, $v_1/c=0.975$, $v_2/c=0.852$. The flow lines for this case are shown in Fig. 4. Of particular interest is the fact that some of the lines close on themselves, indicating that stored energy in the form of circulating waves exists in these regions.

## IV. Conclusions

It has been shown that an exact procedure exists for the design of a modulated corrugated surface to support a specified group of surface waves. In the illustrative examples of the application of this procedure, it has been shown that the surface-impedance level on such a surface attains values well in excess of the corresponding values for the single-mode surfaces designed for the fastest wave in the group. This supports the point of view that a modulated surface can contain a wave with slightly less than the velocity of light as a part of a relatively tightly bound wave system.

The analysis used is limited in application to two-dimensional corrugated surface radiators. It is further limited in that it provides no information on the bandwidth characteristics of the structures designed.

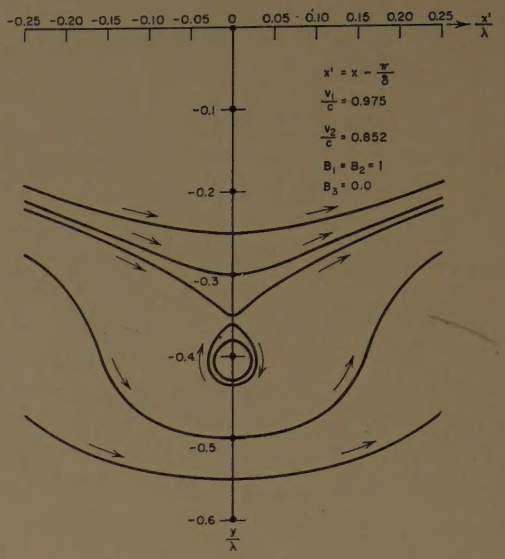



Fig. 4—Energy-flow pattern in vicinity of region of energy reversal.

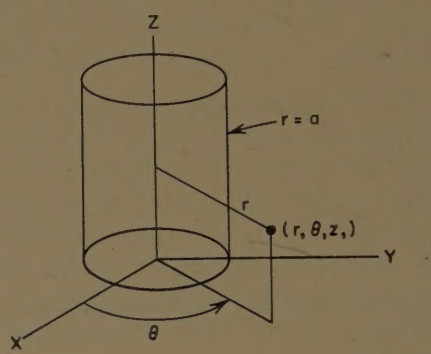



Fig. 5—Cylindrical coordinates system.

## Appendix

### Cylindrical Radiators

The end-fire type of surface wave on a cylindrical surface is described by a combination of the first-order TE- and the first-order TM-cylindrical wave functions,[5]

$$E_r = -jha_1\frac{\partial\psi_1}{\partial r} - \frac{\omega\mu}{r}b_1\psi_1$$

$$E_\theta = -\frac{h}{r}a_1\psi_1 + j\omega\mu b_1\frac{\partial\psi_1}{\partial r}$$

$$E_z = (k^2 - h^2)a_1\psi_1$$

$$H_r = \frac{k^2}{\omega\mu r}a_1\psi_1 - jhb_1\frac{\partial\psi_1}{\partial r}$$

$$H_\theta = \frac{-jk^2}{\omega\mu}a_1\frac{\partial\psi_1}{\partial r} - \frac{h}{r}b_1\psi_1$$

$$H_z = (k^2 - h^2)b_1\psi_1, \qquad (11)$$

[5] J. A. Stratton, "Electromagnetic Theory," McGraw-Hill Book Co., Inc., New York, N. Y.; 1941.

where

$$\psi_1 = \cos\theta H_1^{(2)}(\sqrt{k^2 - h^2}\, r)e^{-jhz}.$$

To calculate the wave constants appropriate to a given cylindrical corrugated surface (it is assumed that there are many corrugations per surface wavelength), one first chooses a relationship between $a_1$ and $b_1$ such that $E_\theta = 0$ at $r = a$, and then selects the value of $h$ which causes $E_z/H_\theta$ to reduce to $jX_s$, where $X_s$ is the surface reactance.

In principle, it is possible to employ two or more cylindrical surface-wave modes having different $a$ and $b$ constants and different values of $h$ in the same way that the planar modes were employed in this paper. Thus, using the analogies $E_y \rightarrow E_r$, $E_x \rightarrow E_z$, and $H_z \rightarrow H_\theta$, it should be possible to define surfaces analogous to those developed here. When this procedure is carried out, however, it is found that the boundary condition $E_\theta = 0$ on the surface is not satisfied in general by the wave groups used to generate the surface function, and hence, the synthesis procedure is not applicable in this case. This difficulty would not arise in the axially-symmetric TM solutions, and an analogous formulation could be developed to provide the characteristics of multimode cylindrical structures for such modes. The axially-symmetric TM modes are of no interest for end-fire antenna applications, however, since they produce radiation-pattern nulls in the axial direction.

# Magneto-Ionic Faraday Rotation of the Radio Signals on 40 Mc from Satellite 1957$_\alpha$ (Sputnik I)*

E. V. SØRENSEN†, ASSOCIATE MEMBER, IRE

*Summary*—**Some of the fadings of the 40-Mc signals from Sputnik I during morning transits are assumed to originate in Faraday rotation in the ionosphere. This has been investigated for the transit of October 12, 1957, 5:30 GMT, for which the orbit was known with reasonable accuracy.**

**The fading records show an increasing fading rate during the last part of the transit. This is contrary to the simple theory based on the assumptions of a flat earth and a homogeneous ionosphere, from which a constant fading rate should be expected. The main reason why simple theory did not apply in this case, is that the satellite travels through the twilight zone from night towards day, *i.e.* from low to high ionization.**

**The Faraday rotation is calculated by means of data for the orbit, magnetic field maps and predictions of the ionosphere conditions. The result is in good agreement with the observation when a high ionization between the altitude of maximum ionization and the satellite is assumed. The effect of refraction is estimated by means of the perturbation theory. It increases the Faraday rotation from 0 to 19 per cent during the transit. The general conclusion is that, even though the fading behavior of a transit near sunrise is explainable to some extent, it is still difficult to deduce detailed information on the state of the ionosphere from it, especially when the frequency is so low that refraction becomes significant.**

* Received by the PGAP, May 19, 1960; revised manuscript received, September 20, 1960.

† Microwave Lab. of the Danish Academy of Tech. Sciences, Copenhagen.

## I. Introduction

IT has been shown[1] that a rather simple interpretation of the fading rate due to magneto-ionic Faraday rotation of plane-polarized radio waves emitted from a satellite is possible when the following approximations apply simultaneously.

1) The curvature of earth can be neglected in the geometry.
2) The satellite travels at a constant height with a constant velocity.
3) The height-profile of the electron density is independent of the geographic position between receiver and satellite.
4) The magnetic field is homogeneous.
5) The signal frequency is so high that refraction is negligible.
6) The quasi-longitudinal condition is fulfilled, *i.e.* the angle between the magnetic field and transmission path differs from 90 degrees by at least a few degrees.

[1] S. A. Bowhill, "The Faraday-rotation rate of a satellite radio signal," *J. Atmos. and Terrest. Phys.*, vol. 13, pp. 175–176; December, 1958.

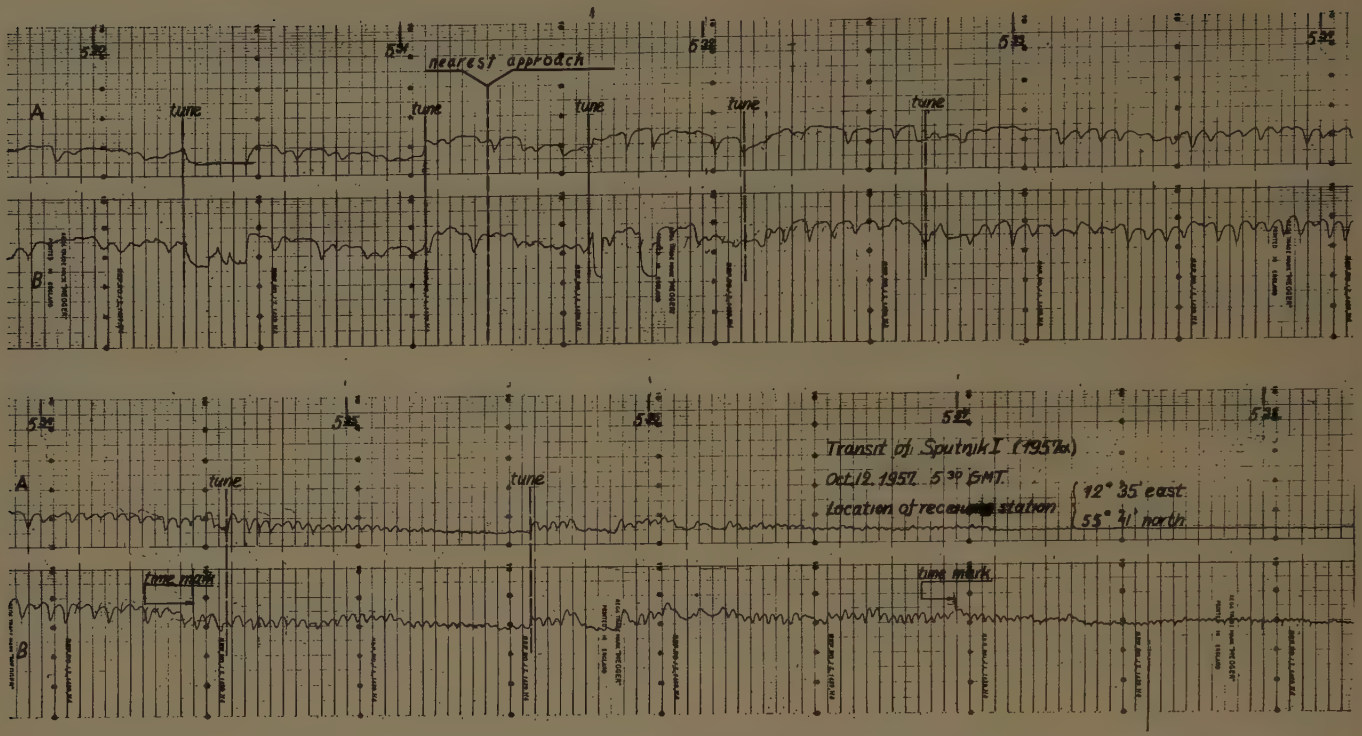



Fig. 1—Fading record for the transit October 12, 1957, 5:30 GMT. Trace *A*, Horizontal dipole. Trace *B*, Vertical monopole. Bandwidth of receivers: 145 cps. Amplitude response: nearly logarithmic.

The fading rate will then be constant in time, independent of receiving site and proportional to:

1) the total amount of electrons in a vertical column below the satellite,
2) the magnitude of the magnetic field component along the satellite orbit,
3) the ratio between velocity and height of the satellite.

For some satellite transits the fades are satisfactorily explained by this simple picture. In other cases, where not all the approximations apply simultaneously, the fading behaviour is more complicated, although the fades in all probability originate in Faraday rotation.

Some of the recorded transits of the first Russian satellite 1957$\alpha$ (Sputnik I) launched October 5, 1957 belong to the last category and have given rise to some discussion.[2] Such a transit will be analyzed in the following. The Faraday rotation will be calculated point by point on the basis of available data for the orbit, the magnetic field and the ionosphere.

The only starting approximation is the assumption of quasi-longitudinal conditions. This will be justified later. The calculated values will be corrected approximately for the effect of refraction. The results agree fairly well with the observations for large values of the electron density between the height of maximum ionization and the satellite. Especially the approximations (3) and (5) have proved to be serious over-simplifications in the present case.

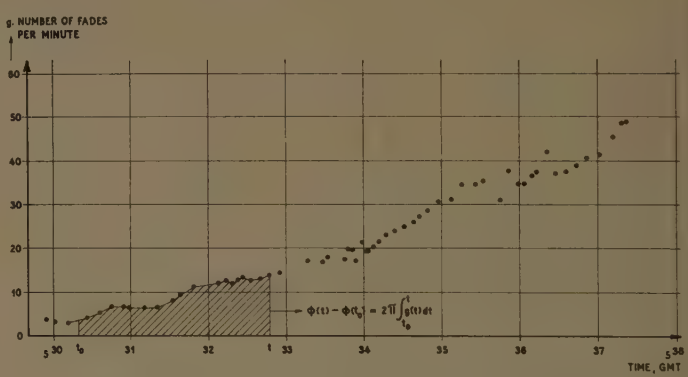



Fig. 2—Number of fades per minute vs time.

Blackband, *et al.*[3] have demonstrated how absolute information on horizontal gradients in the ionosphere may be obtained from such transits. This requires that the point corresponding to quasi-transversal propagation can be located on the fading records, a condition which is not always fulfilled.

## II. Observations

An investigation of the fading characteristics of the 40-Mc radio signals from 1957$\alpha$ (Sputnik I) as received in Copenhagen revealed the following:

1) During the early transits in the evening and the first part of the night, when the satellite passed in alti-

[2] "Radio observations of the Russian satellites, a discussion meeting," *Proc. IEE*, vol. 105, pp. 81–115 (particularly pp. 101–104 and 113); March, 1958.

[3] V. T. Blackband, B. Burgess, I. L. Jones, and G. J. Lawson, "Deduction of ionospheric electron content from the Faraday fading of signals from artificial earth satellites," *Nature*, vol. 183, pp. 1172–1174; April, 1959.

tudes of about 200 kilometers and in directions from south-west to north-east, the signals faded with the constant rates of 6.8 and 13.6 fades per minute independent of the satellite position. These fades are assumed to originate in the satellite spin, and accordingly they are dismissed as being unimportant from the viewpoint of propagation effects.

2) The late transits in the last part of the night and in the morning, when the satellite passed in altitudes between 400 and 500 kilometers and in directions from north-west to south-east, showed fades of gradually increasing rate. In general, the fading started about the time of nearest approach with a rate of some 5 fades per minute and lasted until the signals disappeared, at which time the rate was about 50 fades per minute. Simultaneous recording of the signals in a vertical monopole and a horizontal dipole revealed that the fades occurred nearly in time quadrature on these antennas, as was to be expected for the Faraday rotation type of fading, when the transit takes place near the horizon such as here.

Fig. 1 shows such a record of the latter half part of the transit October 12, 1957, 5:30 GMT (the lines marked "tune" indicate that the common local oscillator for the two receivers had been adjusted manually to compensate for the Doppler-frequency shift). In Fig. 2 the Faraday rotation fading rate, as found by an examination of the record, is plotted against time. Sudden jumps amounting to one or two times the RPM of the satellite spin may occur in such a curve. In the present case however, the fading rate varies in a fairly continuous way, indicating that the fades almost certainly originate in Faraday rotation.

Fig. 3 shows the corresponding part of the orbit as established from scattered information. The different points with the corresponding indications of time have been used in the calculations of the Faraday rotation.

If the assumption that these fades originate in Faraday rotation is correct, the variation with time of the phase difference between the ordinary and the extraordinary ray may be obtained from graphical integration of the fading rate $g(t)$ as indicated in Fig. 2.

$$\phi(t) - \phi(t_0) = 2\pi \int_{t_0}^{t} g(t)dt \quad \text{rad.} \qquad (1)$$

If we choose $t_0$ as the time when the ray path is perpendicular to the magnetic field, $\phi(t_0)$ will be very close to zero. The corresponding satellite position is somewhat to the north of point $A$ in Fig. 3. This time cannot, however, be found from the fading record as it corresponds to an infinitely long fading period. Accordingly we have chosen the time (5:30.19 GMT) from which the fade can be traced with reasonable certainty. This corresponds to point $A$ in Fig. 3, for which the angle between the ray path and the magnetic field is 67 degrees. The result of the integration is plotted in Fig. 4 together with the calculated curve.

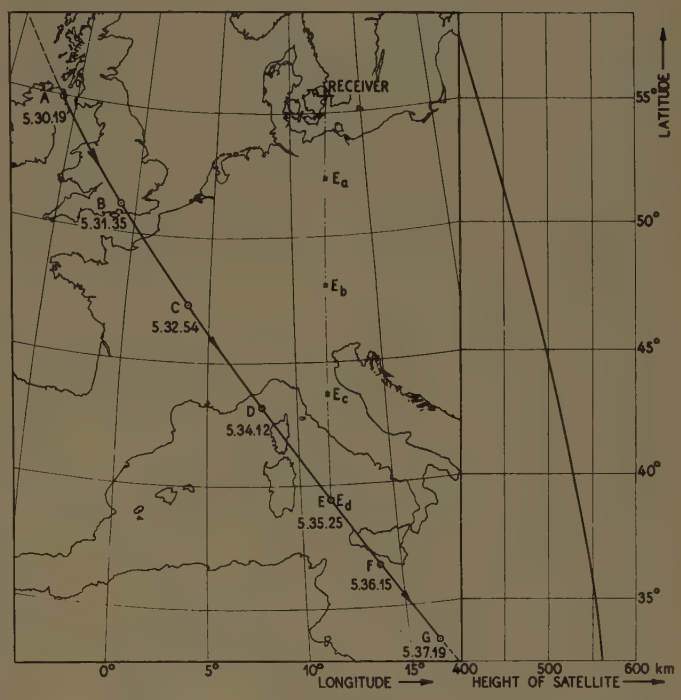



Fig. 3—Assumed orbit for 1957$\alpha$ during transit October 12, 1957, 5:30 GMT.

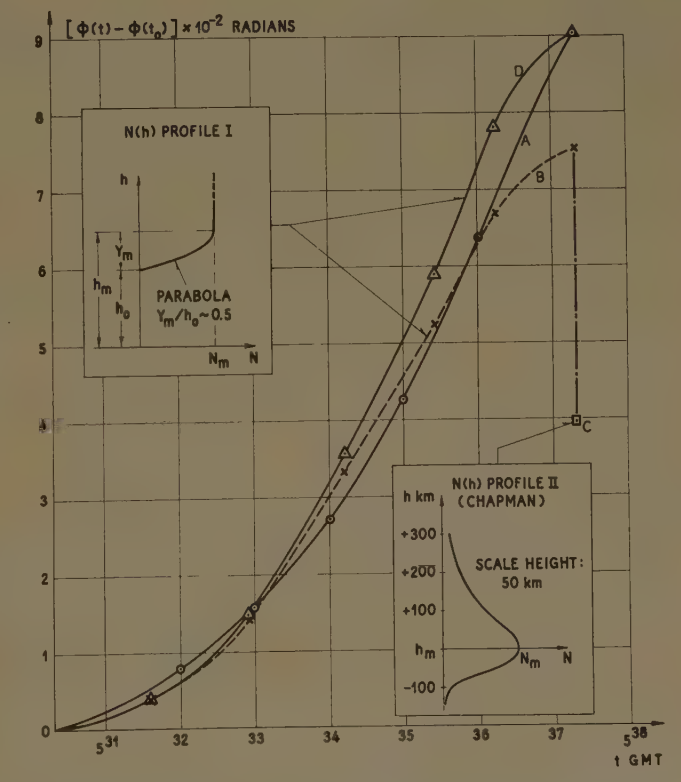



Fig. 4—Variation with time of the phase difference $\phi$ between ordinary and extraordinary rays. Curve $A$, obtained from observed fading rate. Curve $B$ and point $C$, calculated by means of predicted data for the magnetic field and for the ionosphere. Different $N(h)$ profiles. Curve $D$, calculated curve corrected for the effect of refraction.

## III. Theory

As shown by Appleton in 1932[4] the refractive index of the ionosphere, taking the magnetic field of earth into consideration is given by:

$$n_{1,2} = \sqrt{1 - \frac{2}{2\alpha - \frac{\gamma_T^2}{\alpha - 1} \pm \sqrt{\frac{\gamma_T^4}{(\alpha-1)^2} + 4\gamma_L^2}}} \quad (2)$$

where:

$$\alpha = \frac{m\omega^2}{4\pi Ne^2}; \quad \gamma_L = \frac{\omega H \cos\theta}{4\pi Nec} \quad \text{and} \quad \gamma_T = \frac{\omega H \sin\theta}{4\pi Nec}. \quad (3)$$

$H$ is the local magnetic field in oersteds
$\theta$ is the local angle between the ray path and the magnetic field
$N$ is the local ionization in electrons per cm$^3$
$e$ is the charge of an electron: $-4.8\cdot10^{-10}$ esu
$m$ is the mass of an electron: $9.11\cdot10^{-28}$ g
$c$ is the velocity of light: $3\times10^{10}$ cm/sec

If the quasi-longitudinal condition holds, (2) may be reduced to

$$n_{1,2} \sim 1 - \frac{1}{2\alpha} \pm \frac{\gamma_L}{2\alpha^2} = n_0 \pm \frac{1}{2}\Delta n, \quad (4)$$

where $n_0$ is independent to the magnetic field.

In the present analysis, the "worst" value of $\theta$ occurs in point $A$ Fig. 3 ($\theta = 67$ degrees). The magnetic field is here 0.4 oersted. If an ionization of $N = 4\cdot10^5$ electrons per cm$^3$ is assumed, and the frequency is 40 Mc, we find:

| | Calculated from (2) | Calculated from (4) |
|---|---|---|
| $1-n_2$ | $1.024\cdot10^{-2}$ | $1.018\cdot10^{-2}$ |
| $1-n_1$ | $1.002\cdot10^{-2}$ | $0.996\cdot10^{-2}$ |
| $\Delta n = n_1 - n_2$ | $2.196\cdot10^{-4}$ | $2.171\cdot10^{-4}$ |

which indicates that (4) is a reasonable approximation in the following analysis.

Under the same condition the two magneto-ionic components of the radio wave from the satellite will be nearly circularly polarized TEM fields rotating in opposite directions. When they arrive at the receiver, the phase difference between them will be:

$$\phi = \frac{\omega}{c}\left[\int_S^R n_1 ds_1 - \int_S^R n_2 ds_2\right] \text{rad}. \quad (5)$$

As the satellite moves, $\phi$ will vary with time. Each alteration amounting to $2\pi$ will turn the plane of polarization of the resulting received signal by $\pi$ radians and produce one fade. The function $\phi(t)$ is determined from calculation of $\phi$ for a suitable number of points of the orbit (points $A$–$G$, Fig. 3). From this the quantity $\phi(t) - \phi(t_0)$ may be compared to the values found from the graphical integration outlined in Section II.

A rigorous calculation of the integrals in (5) is out of the question as it requires the tracing of the two rays through a medium with an inhomogeneous magnetic field and electron distribution. An approximate solution may assume the following form, where (5) is rewritten to:

$$\phi \sim F\cdot\frac{\omega}{c}\int_S^R (n_1 - n_2)ds = F\cdot\frac{\omega}{c}\int_S^R \Delta n ds. \quad (6)$$

Here the integral is taken along the optical path between $S$ and $R$, and the refraction is taken into consideration through the approximate correction factor $F$ (Section V and Appendix).

Substitution of (3) and (4) in (6) gives

$$\phi \sim F\cdot\frac{e^3}{\pi\cdot c^2\cdot m^2\cdot f^2}\int_S^R N(s)\cdot H(s)\cdot\cos\theta\,(s)ds. \quad (7)$$

The integral in (7) is convenient for graphical solution when the variations of $N$, $H$ and $\cos\theta$ along the optical path are known. The establishment of the electron density data, for which there is considerable uncertainty, will be outlined in the following. The mapping of the two other quantities rests upon charts of the geomagnetic field[5] and the following information on the orbit:

1) The inclination of the orbit is about 65 degrees.
2) The transit in question is an overhead transit for the Royal Aircraft Establishment radio station in Lasham, England.[6] Incidentally this point is almost coincident with transit point as seen from Copenhagen.
3) The vertical motion of the satellite was derived from a curve giving height and latitude vs the time since the preceding northward equatorial transit.[7]

## IV. The Ionosphere Condition

The first assumption in establishing the $N(s)$ data is that only the $F_2$ layer is present. This is very probable in view of the hour of the day.

The maximum electron density $N_m$ and the corresponding altitude $h_m$ are mapped by means of charts of the predicted $f_0F_2$—and MUF (4000 km)—frequencies for October 1957, (Slough Bulletin A—No. 128), in con-

[4] E. V. Appleton, "Wireless studies of the ionosphere," *J. IEE*, vol. 71, pp. 642–650; October, 1932.

[5] E. H. Vestine, L. Laporte, C. Cooper, I. Lange, and W. C. Hendrix, "The Description of the Earth's Main Magnetic Field and Its Secular Change, 1905–1945," Carnegie Inst., Washington, D. C., No. 578; 1947.

[6] Radio observations of the Russian satellites, a discussion meeting, *op. cit.*, p. 94.

[7] H. Uyeda, T. Ishida, H. Shibata, and M. Mambo, "On the reception of radio waves from Russian Earth Satellite I," *J. RRL* (Japan), vol. 20, pp. 135–141; April, 1958.

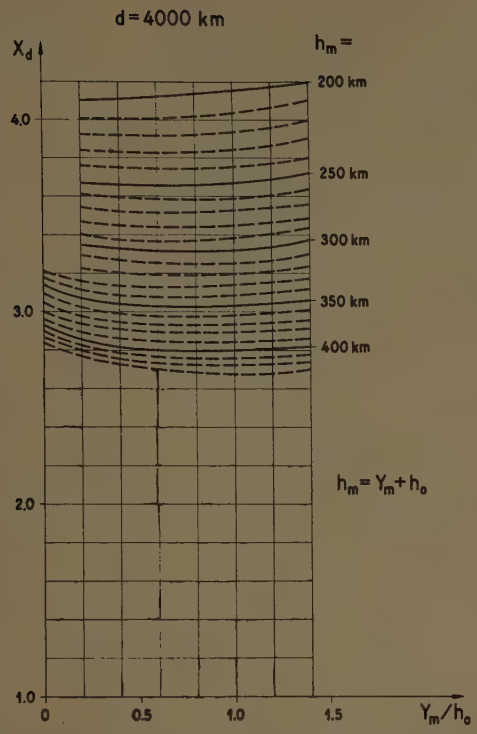



Fig. 5—Dependence of the MUF factor $X_d$ on height $h_m$ of maximum electron density and semi-thickness $Y_m$ of a parabolic $F_2$ layer (Appleton and Beynon[8]).

nection with Appleton-Beynon's theory on MUF calculation.[8] $N_m$ is given by $f_0F_2$:

$$N_m = 1.24 \cdot 10^4 (f_0F_2 \text{ Mc})^2 \ e/\text{cm}^3 \qquad (8)$$

$h_m$ is found by means of the MUF-factor defined by

$$x_d = \frac{\text{MUF}(d)}{f_0F_2} \cdot \qquad (9)$$

Fig. 5 (taken from Appleton and Beynon[8]) shows that this quantity is a sensitive function of $h_m$, whereas it is fairly independent of $y_m/h_0$ where $y_m$ is the semi-thickness of the parabolic $N(h)$ profile assumed in the theory, and $h_0$ is the height of the lower boundary above the ground. The values thus found of $N_m$ and $h_m$ below the satellite and above Copenhagen during the fading-period of the transit are shown in Fig. 6. The motion of the satellite through the twilight zone from night towards day explains the increasing value of $N_m$, and this again is the principal reason for the increasing fading rate characterizing the morning transits of this satellite as observed from Copenhagen.

The knowledge of $N_m$ and $h_m$ is not sufficient to obtain $N(s)$. We still need to form an idea of the remaining part of the $N(h)$ profile, and here the greatest amount of uncertainty is met. According to Appleton

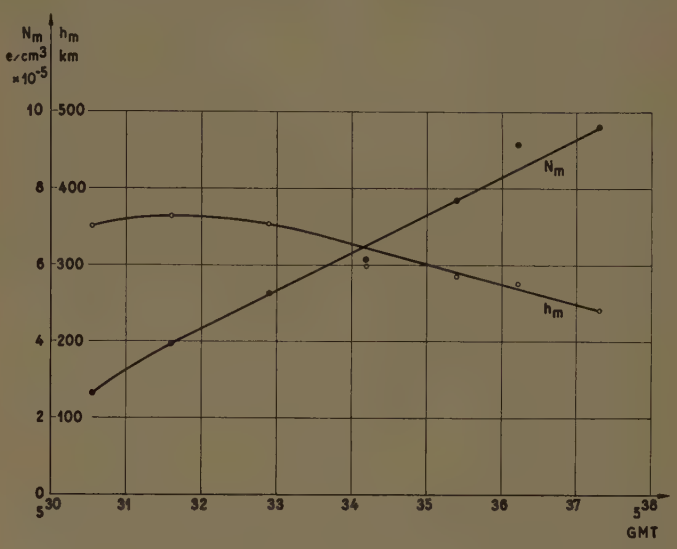



Fig. 6—Variation with time of $N_m$ and $h_m$ below satellite, as derived from ionosphere predictions for October 1957. The values over Copenhagen during the transit were: $N_m \sim 4.3$–$4.6 \times 10^5$ $e/\text{cm}^3$, $h_m \sim 300$ km.

and Beynon[9] a fairly good approximation to the $N(h)$ profile below $h_m$ is a parabola with $y_m/h_0 \sim 0.5$.

We know as yet very little about the electron distribution above $h_m$. An extension of the parabola above $h_m$ gives too few electrons to account for the magnitude of Faraday rotation observed. This is also the case if we make use of a Chapman distribution[10] with usual values of the scale height.[11] This is consistent with the few rocket measurements of the upper ionosphere reported so far[12] and with recent theories of whistler propagation[13] and the generation of low frequency signals in the exosphere,[14] which all indicate that the electron density above $h_m$ decreases more slowly with the height than previously assumed. In view of the relatively low satellite height we therefore make the convenient assumption that the electron density is constant and equal to $N_m$ between $h_m$ and the satellite. As shown in Fig. 4 this gives a fairly good fit to the observed Faraday rotation. Blackband, *et al.*[3] have used the same approximation in their work, and their results proved to be consistent with vertical incidence soundings. Fig. 7 shows

[8] E. V. Appleton and W. J. G. Beynon, "The application of ionospheric data to radio communication problems: Part I," *Proc. Phys. Soc.*, vol. 52, pp. 518–533; July, 1940; "The application of ionospheric data to radio communication problems: Part II," *Proc. Phys. Soc.*, vol. 59, pp. 58–76; January, 1947.

[9] Appleton and Beynon, *op. cit.*, vol. 59, p. 65.

[10] S. K. Mitra, "The Upper Atmosphere," The Asiatic Soc. of Calcutta, India; 1952.

[11] The result of $[\phi(t) - \phi(t_0)] \times 1/F$ where $t$ corresponds to point $G$. Fig. 3 is shown in Fig. 4 for the case of a Chapman distribution with scale height $H = 50$ km. The distribution is given by

$$N_h = N_m e^{1/2(1-z} - e^{-z)}$$

where

$$z = \frac{h - h_m}{H} \cdot$$

[12] V. I. Krassovsky, "Exploration of the upper atmosphere with the help of the third Soviet Sputnik," PROC. IRE, vol. 47, pp. 289–296; February, 1959.

[13] R. A. Helliwell and M. G. Morgan, "Atmospheric whistlers," PROC. IRE, vol. 47, pp. 200–208; February, 1959.

[14] R. M. Gallet, "The very-low-frequency emissions generated in the earth's exosphere," PROC. IRE, vol. 47, pp. 211–231; February, 1959.

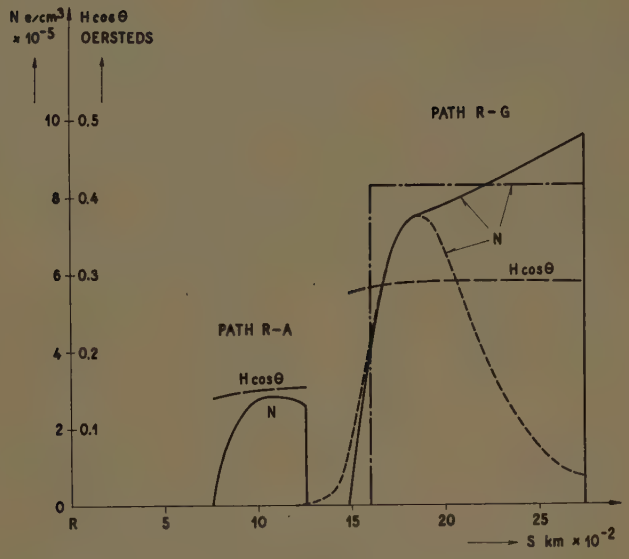



Fig. 7—Examples of the functions $N(s)$ and $H(s)\cdot\cos\theta(s)$.
———— $N_1(s)$ according to $N(h)$—profile I (Fig. 4)
– – – – – – $N_2(s)$ according to $N(h)$—profile II (Fig. 4)
——— – ——— Approximation to $N_1(s)$ for calculation of refraction.

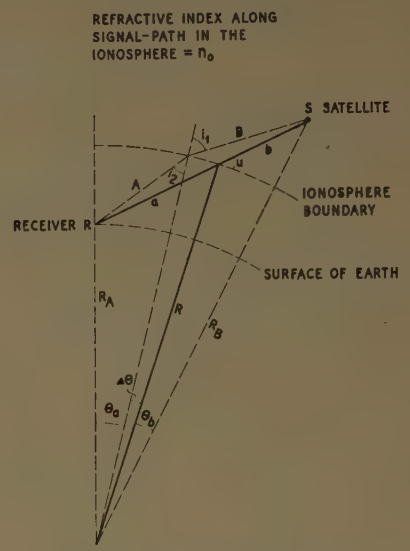



Fig. 8—Geometry involved in calculation of refraction correction factor.

the variation of $N$ and $H\cos\theta$ along some of the integration paths. The last quantity is remarkably constant along each path, which means that the Faraday rotation is a measure of the amount of electrons between receiver and satellite.

## V. Effect of Refraction

A perturbation calculation of the refraction factor $F$ in (6) is given in the Appendix. The underlying assumptions are:

1) The refraction is so small that the values of $N$, $H$ and $\cos\theta$ along the optical path can be used.
2) The product $NH\cos\theta$ is constant along the ionospheric part of the optical path. The validity of this approximation may be judged from Fig. 7 path *R–G*.
3) The tilt of the ionosphere boundary relative to earth is negligible around the penetration-point of the ray.

The refractive index of the ionosphere along one of the transmission paths under consideration will now be a constant depending only on the satellite position. The result is

$$F \sim 1 + (1 - n_0)\frac{a}{a+b}\cot^2 u \qquad (10)$$

where $n_0$ is defined by (4); $a$ and $b$ are the lower and upper parts respectively of the optical path as divided by the ionosphere boundary, and $u$ is the inclination of the optical path at the penetration point (see Fig. 8). $F$ was found to increase from approximately 1 to 1.19 during the transit. The corrected curve for $(\phi(t)-\phi(t_0))$ is shown in Fig. 4.

## VI. Comments

As may be seen from Fig. 4, the calculated total number of fades in the period under consideration agree very closely with the observed number. The slope of the curves, based upon calculation or observation, reflecting the instantaneous fading rate differ somewhat more, especially towards the end of the period. This may be due to local deviations from the predicted ionosphere conditions.

The results indicate a relatively high ionization above the height of maximum ionization.

The effect of refraction is far from negligible even at the relatively high frequency of 40 Mc. In our case it increases the total number of fades by 19 per cent.

The instantaneous fading rate cannot be taken as a measure of the electron content in a vertical column below the satellite (as predicted by simply theory) when the ionization varies along the orbit. If the refraction can be neglected, it is still a measure of the number of electrons between the satellite and the observer. However, this condition is not sufficiently fulfilled at 40 Mc.

In view of this examination we conclude that detailed information on the ionosphere can hardly be expected from fading records corresponding to transits near sunrise, unless higher frequencies are used or more refined and sensitive methods of analysis are developed.

Finally, reference should be made to additional theoretical and observational work on this subject. Aitchison and Weekes[15] have gained ionosphere data from Faraday rotation rate measurements at 20 and 40 Mc for 1957$\alpha$2. Their analysis includes a vertical component of the satellite motion, and the effect of refraction is

[15] G. J. Aitchison and K. Weekes, "Some deductions of ionospheric information from the observations of emissions from satellite 1957 $\alpha$2.-I," *J. Atmos and Terrest. Phys.*, vol. 14, pts. 3 and 4, pp. 236–243; 1959.

estimated. They use a horizontally stratified ionosphere model and exclude transits near sunrise.

Garriott[16] has used two techniques to obtain electron density data from radio observations of Sputnik III during a period of eight months. One method involves a determination of the total rotation angle at the transit point. The other is based on rotation rate measurements. A complete account of all the approximations is given and the relevant corrections are made. Horizontal gradients appear to be included in their computation program.

## Appendix

## Calculation of the Effect of Refraction

The geometry of the problems is shown in Fig. 8. The ionosphere-approximation applied in this model is explained in Section V.

### *General Method*

First the magnetic field is neglected and the electrical distance $\Psi$ in radians between satellite and receiver is expanded in terms of the variable: $x=1-n_0$:

The magnetic field then is introduced as a perturbation, and the phase difference between the two resulting rays is calculated as

$$\phi = \Delta\Psi = -\frac{\delta\Psi(x)}{\delta x}\Delta n. \tag{11}$$

The electrical length of the refracted ray path $A-B$ in Fig. 8 is

$$\Psi = \frac{\omega}{c}(A + B(1 - x)). \tag{12}$$

The ray segments $A$ and $B$ may be expanded in terms of the infinitesimal angle $\Delta\theta$ by means of simple geometrical relations:

$$A = a - L\Delta\theta + M\Delta\theta^2 + \cdots \tag{13}$$

$$B = b + L\Delta\theta + N\Delta\theta^2 + \cdots. \tag{14}$$

$\Delta\theta$ in turn may be expanded in terms of $x=1-n_0$ by means of the law of refraction:

$$\Delta\theta = Px + Qx^2 + \cdots \tag{15}$$

substitution of (12)–(15) in (11) gives

$$\phi = \left[1 + 2\frac{LP - (M+N)P^2}{b}x + \cdots\right]\frac{\omega}{c}b\Delta n. \tag{16}$$

The correction factor $F$ in (6) is now recognized as the term in the brackets. Note that the first order approximation to $\phi$ requires both the linear and the quadratic terms of (13) and (14), but only the linear term of (15) as $Q$ is absent in the coefficient to $x$. This is a considerable simplification in the detailed calculations. The result of the somewhat comprehensive procedure is (referring to Fig. 8):

$$\phi \sim \left[1 + (1 - n_0)\frac{a}{a+b}\cot^2 u\right]\frac{\omega}{c}b\Delta N. \tag{17}$$

Note that the curvature of earth is absent in this approximation.

## Acknowledgment

The author is indebted to the whole staff of the Microwave Laboratory for their prompt effort to record the signals from 1957$\alpha$ immediately after it had been launched. M. Grønlund has contributed to the analysis with many valuable suggestions, especially with regard to the refraction problem. The help and advice of E. Ungstrup concerning the ionosphere data are also gratefully acknowledged.

[16] O. K. Garriott, "The determination of ionospheric electron content and distribution from satellite observations," *J. Geophys. Res.*, vol. 65, pp. 1139–1158; April, 1960.

# On the Guided Propagation of Electromagnetic Wave Beams*

G. GOUBAU†, FELLOW, IRE, AND F. SCHWERING†, MEMBER, IRE

*Summary*—**Any field in a half-space can be described by a continuous spectrum of cylindrical waves. If this spectrum comprises substantially only waves whose propagation constant is very close to the plane wave propagation constant, the field can be resolved into a set of elementary wave beams which are characterized by Laguerre polynomials. They satisfy orthogonality relations like the wave modes in a waveguide. The elementary beams or "beam modes" can be reiterated and guided by reconstituting the cross-sectional phase distribution at certain intervals. Reiterative beams are utilized in the beam waveguide. The finite size of the phase resetting devices effects a modification of the reiterative beam modes and causes diffraction losses. These losses decrease very rapidly with increasing diameter of the phasing devices.**

## Introduction

THE wave beam emitted by a highly directional antenna, such as a parabolic antenna, has within the Fresnel region a substantially uniform diameter. However, the cross-sectional amplitude and phase distributions vary substantially along the path of the beam.[1] These variations depend on the field distribution at the antenna. The question may be asked whether there are beams in which the cross-sectional amplitude distribution is repeated at a certain distance from the origin of the beam. If there are any such beams, the original phase distribution can be reconstituted by suitable phase shifting means so that a new Fresnel region is formed which has the same field configuration as the original one. Repeating this iteration process by passing the beam along a structure of uniformly spaced phasing devices (phase transformers), the beam is guided without expansion of energy.

The main objective of this paper is to show that there are indeed reiterative wave beams. A further objective is the calculation of the diffraction loss caused by the finite dimensions of the phase transformers.

## Reiterative Wave Beams

We consider wave beams which propagate along the $z$-axis of a cylindrical coordinate system, Fig. 1. The general solution of Maxwell's equations for a homogeneous half-space free of sources can be constructed of partial solutions with the components[2]

$$E_\rho = E_{\rho m}^{(\pm)} \cos(m\phi + \alpha_m)$$

$$E_\phi = E_{\phi m}^{(\pm)} \sin(m\phi + \alpha_m)$$

$$E_z = E_{zm}^{(\pm)} \cos(m\phi + \alpha_m)$$

$$H_\rho = \mp \sqrt{\frac{\epsilon}{\mu}} E_{\rho m}^{(\pm)} \sin(m\phi + \alpha_m)$$

$$H_\phi = \pm \sqrt{\frac{\epsilon}{\mu}} E_{\phi m}^{(\pm)} \cos(m\phi + \alpha_m)$$

$$H_z = \mp \sqrt{\frac{\epsilon}{\mu}} E_{zm}^{(\pm)} \sin(m\phi + \alpha_m), \tag{1a}$$

with

$$E_{\rho m}^{(\pm)} = \frac{1}{2}\int_0^\infty f_m^{(\pm)}(\gamma)\{(k \pm h)J_{m+1}(\gamma\rho) + (k \mp h)J_{m-1}(\gamma\rho)\}e^{-jhz}\gamma d\gamma$$

$$E_{\phi m}^{(\pm)} = \frac{1}{2}\int_0^\infty f_m^{(\pm)}(\gamma)\{(k \pm h)J_{m+1}(\gamma\rho) - (k \mp h)J_{m-1}(\gamma\rho)\}e^{-jhz}\gamma d\gamma$$

$$E_{zm}^{(\pm)} = \mp j\int_0^\infty f_m^{(\pm)}(\gamma)J_m(\gamma\rho)e^{-jhz}\gamma^2 d\gamma. \tag{1b}$$

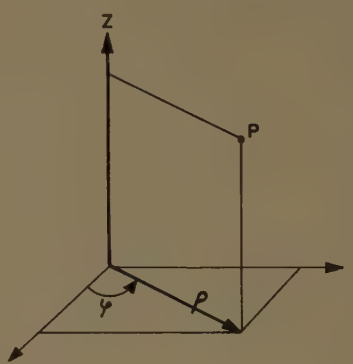



Fig. 1.

$f_m^{(+)}(\gamma)$, $f_m^{(-)}(\gamma)$ are arbitrary functions of the real variable $\gamma$, subject to the condition that the integrals in (1b) and the integrals over the first derivatives of the integrants with regard to $\rho$ and $z$ exist. $f_m^{(+)}(\gamma)$, $f_m^{(-)}(\gamma)$ represent the amplitude spectra of the elementary cylindrical waves which form the beam. The propagation constant $h$ is determined by the relation $k^2=\gamma^2+h^2$, where $k=2\pi/\lambda$ is the plane wave propagation constant in the half space. $J_{m-1}$, $J_m$ and $J_{m+1}$ are the Bessel functions of the order $m-1$, $m$ and $m+1$, respectively. $m$ is any positive integer including zero.

For symmetry reasons, a field representation in terms of hybrid waves has been chosen in preference to the more conventional one in $E$-waves and $H$-waves. There are two alternatives for combining $E$- and $H$-waves of the same order to hybrid waves symmetrical in $E$ and $H$. They are distinguished in (1a) and (1b) by the super-

* Received by the PGAP, June 16, 1960.
† USASRDL, Fort Monmouth, N. J.
[1] S. Silver, "Microwave Antenna Theory and Design," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., p. 173; 1941.
[2] J. A. Stratton, "Electromagnetic Theory," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 360–371; 1951.

scripts (+) or (−). Pure $E$- or $H$-wave fields are obtained by superposition of two hybrid fields with $f_m^{(+)}(\gamma) = \mp f_m^{(-)}(\gamma)$.

The elementary waves are propagating waves in the range $0<\gamma<k$ and evanescent waves in the range $\gamma<k<\infty$.

Assume the beam contains only propagating waves and $f_m^{(+)}(\gamma)$ and $f_m^{(-)}(\gamma)$ are real functions. Then, as evidenced by (1a) and (1b), the transverse field components at $z=+z_0$ and $z=-z_0$ are conjugate complex. Assume, furthermore, that there is a phase transformer which, when inserted into the beam at $z=z_0$, transforms the phase distribution at this plane into a phase distribution identical with that at $z=-z_0$. Then, the original beam is reconstituted; this means, the field distribution in the space range $-z_0<z<+z_0$ is repeated in the range $+z_0<z<+3z_0$. If several of these phase transformers are placed at regular intervals of $2z_0$, a reiterative beam is obtained, which is "guided" by this phase shifting structure.

The question as to the physical realizability of the phase transformers depends on the structure of the beam. Fortunately, there are beams where the required phase transformation is obtainable with very simple means, at least within a finite cross-sectional area. These beams are characterized by a $\gamma$-spectrum substantially limited to values of $\gamma^2 \ll k^2$. In other words, the functions $f_m^{(\pm)}(\gamma)$ have appreciable values only within a range $0<\gamma<\gamma_c$ where $\gamma_c^2 \ll k^2$. In this case, the propagation constant $h$ in the phase terms of (1b) can be approximated by $h \approx k-\gamma^2/2k$, provided $(\gamma_c^2/2k)z_0$ remains within the order of $2\pi$. This imposes a limitation of the $z_0$-range in which the approximation is valid. In the amplitude terms of the transverse field components, (1b), $h$ can be replaced by $k$. Thus, (1b) simplifies to

$$E_{\rho m}^{(\pm)} = \pm E_{\phi m}^{(\pm)} = ke^{-jkz}\int_0^\infty f_m^{(\pm)}(\gamma)J_{m\pm1}(\gamma\rho)e^{+j(\gamma^2/2k)z}\gamma d\gamma$$

$$E_{zm}^{(\pm)} = \mp jke^{-jkz}\int_0^\infty f_m^{(\pm)}(\gamma)J_m(\gamma\rho)e^{+j(\gamma^2/2k)z}\frac{\gamma^2}{k}d\gamma. \quad (2)$$

Since the field is uniquely determined by the transverse field components, only these components are considered in the following.

Due to the restriction on $f_m^{(\pm)}(\gamma)$, the transverse electric as well as the magnetic field components have the same cross-sectional distribution. The phase transformation which is required for reconstituting or iterating such a beam is indeed feasible, at least theoretically, as shown in Appendix I.

Forthcoming deductions prove it very advantageous to represent $f_m^{(\pm)}(\gamma)$ in (2) by the following series of orthogonal functions:

$$f_m^{(\pm)}(\gamma) = \sum_{n=0}^{\infty}\alpha_{n,m}^{(\pm)}f_{m,n}^{(\pm)}(\gamma) = \sum_{n=0}^{\infty}\alpha_{n,m}^{(\pm)}\left(\frac{\gamma}{\gamma_0}\right)^{m\pm1} \cdot L_{n,m\pm1}\left(\frac{\gamma^2}{\gamma_0{}^2}\right)e^{-1/2(\gamma^2/\gamma_0{}^2)}, \quad (3)$$

where $\gamma_0$ is an arbitrary constant.

$$L_{n,l}(x) = \sum_{i=0}^{n}\binom{n+l}{n-i}\frac{(-x)^i}{i!}$$

are the generalized Laguerre polynomials,[3] and

$$\alpha_{n,m}^{(\pm)} = 2\frac{n!}{(n+m\pm1)!}\int_0^\infty f_m^{(\pm)}(\gamma)\left(\frac{\gamma}{\gamma_0}\right)^{m\pm1+1} \cdot L_{n,m\pm1}\left(\frac{\gamma^2}{\gamma_0{}^2}\right)e^{-1/2(\gamma^2/\gamma_0{}^2)}d\left(\frac{\gamma}{\gamma_0}\right)$$

are the coefficients of expansion. The Laguerre polynomials $L_{n,l}(x)$ form for each $l$ a complete system of orthogonal functions for the range $0<x<\infty$ with the weight function $x^le^{-x}$.

The fields derived from $f_0^{(+)}(\gamma)$ and $f_0^{(-)}(\gamma)$ are linearly dependent, since the field components, [(1a), (1b)], with the superscripts (+) and (−), transform into each other if $\alpha_0$ is replaced by $\pi-\alpha_0$. Therefore, only one of the functions, namely $f_0^{(+)}$, is considered in the following.

Now, (2) can be evaluated, and we obtain the field distribution functions[4]

$$\begin{aligned}E_{\rho m}^{(\pm)} &= \pm E_{\phi m}^{(\pm)} \\ &= ke^{-jkz}\sum_{n=0}^{\infty}\alpha_{n,m}^{(\pm)}(-1)^n\frac{\gamma_0{}^{m\pm1+2}}{\left(1-j\frac{\gamma_0{}^2}{k}z\right)^{m\pm1+1}} \\ &\cdot e^{j2n\,\mathrm{arctg}\,[(\gamma_0{}^2/k)z)]}\rho^{m\pm1}L_{n,m\pm1}\left[\frac{\gamma_0{}^2\rho^2}{1+\left(\frac{\gamma_0{}^2}{k}z\right)^2}\right] \\ &\cdot e^{-1/2(\gamma_0{}^2\rho^2)/[1-j(\gamma_0{}^2/k)z]}.\end{aligned} \quad (4)$$

The beam can thus be considered as the superposition of elementary beams having the field components

$$\begin{aligned}E_{\rho m,n}^{(\pm)}e^{jkz} &= \pm E_{\phi m,n}^{(\pm)}e^{jkz} \\ &= A_{m,n}^{(\pm)}\frac{u^{m\pm1}}{(1+v^2)^{1/2(m\pm1+1)}}L_{n,m\pm1}\left(\frac{u^2}{1+v^2}\right) \\ &\cdot \exp\left\{-\frac{1}{2}\frac{u^2}{1+v^2}+j\left[(2n+m\pm1+1,\right.\right. \\ &\left.\left.\cdot\,\mathrm{arctg}\,v-\frac{1}{2}\frac{u^2v}{1+v^2}\right]\right\},\end{aligned} \quad (5a)$$

$$\begin{aligned}H_{\rho m,n}^{(\pm)} &= \mp H_{\phi m,n}^{(\pm)} = \mp\sqrt{\frac{\epsilon}{\mu}}E_{\rho m,n}^{(\pm)} \\ &= -\sqrt{\frac{\epsilon}{\mu}}E_{\phi m,n}^{(\pm)}\end{aligned} \quad (5b)$$

[3] W. Magnus and F. Oberhettinger, "Functions of Mathematical Physics," Chelsea Publishing Co., New York, N. Y., pp. 84–86; 1955.

[4] A. Erdely, "Tables of Integral Transforms," McGraw-Hill Book Co., Inc., New York, N. Y., vol. 2, p. 43; 1954.

with

$$u = \gamma_0 \rho = \left(\frac{\gamma_0}{k}\right)(k\rho), \quad v = \frac{\gamma_0^2}{k} z = \left(\frac{\gamma_0}{k}\right)^2 (kz),$$

$$A_{m,n}^{(\pm)} = (-1)^n \gamma_0^2 k.$$

For $m=0$, only the upper sign is valid.

The elementary beams or "beam modes" satisfy the same orthogonality relations as the modes in an ordinary loss-less waveguide

$$\int_F \{E_{m,n}^{(\pm)} \times H_{m',n'}^{(\pm)*}\} e_z dF = \sqrt{\frac{\epsilon}{\mu}} A_{m,n}^{(\pm)} \cdot A_{m',n'}^{(\pm)*} \frac{\pi}{\gamma_0^2} \frac{(n+m\pm 1)!}{n!} \delta_{nn'} \cdot \delta_{mm'}$$

$$\int_F \{E_{m,n}^{(\pm)} \times H_{m',n'}^{(\mp)*}\} e_z dF = 0, \tag{6a}$$

$$\int_F E_{m,n}^{(\pm)} E_{m',n'}^{(\pm)*} dF = A_{m,n}^{(\pm)} A_{m',n'}^{(\pm)*} \frac{\pi}{\gamma_0^2} \cdot \frac{(n+m\pm 1)!}{n!} \delta_{nn'} \delta_{mm'}$$

$$\int_F E_{m,n}^{(\pm)} E_{m',n'}^{(\mp)*} dF = 0 \qquad \text{with } \delta_{nn'} = \begin{matrix} 1 & n = n' \\ 0 & n \neq n', \end{matrix} \tag{6b}$$

where the integration is performed over any plane $z=$ constant; $e_z$ is the unit vector of the $z$-direction; $E_{m,n}^{(\pm)}$ and $H_{m,n}^{(\pm)}$ denote the transverse field vectors of the beam modes and $E_{m,n}^{(\pm)*}$, $H_{m,n}^{(\pm)*}$ the conjugate complex vectors. The proof of (6) is given in Appendix II.

Eq. (6a) shows that the total power of a beam comprising several beam modes is the sum of the powers of the individual beam modes. The power of a mode is given by

$$P_{n,m}^{(\pm)} = \frac{1}{2}\sqrt{\frac{\epsilon}{\mu}} \cdot A_{m,n}^{(\pm)} \cdot A_{m,n}^{(\pm)*} \frac{\pi}{\gamma_0^2} \frac{(n+m\pm 1)!}{n!} \qquad \text{for } m \neq 0$$

$$P_{n,0}^{(+)} = \frac{1}{2}\sqrt{\frac{\epsilon}{\mu}} \cdot A_{0,n}^{(+)} \cdot A_{0,n}^{(+)*} \frac{\pi}{\gamma_0^2} \cdot \frac{(n+1)!}{n!} \qquad \text{for } m = 0. \tag{7}$$

The vector functions $E_{m,n}^{(+)}$ and $E_{m,n}^{(-)}$ form a complete system of orthogonal functions. With reference to (1a), we note that each vector function represents *two* modes characterized by $\alpha_m=0$ and $\alpha_m=\pi/2$, respectively. Any given field distribution in a plane $z$-constant can mathematically be developed into the mode functions. However, this development has physical meaning only if the corresponding amplitude distribution functions $f_m^{(\pm)}(\gamma)$ satisfy the condition introduced at the beginning of this paragraph; *viz*, that $f_m^{(\pm)}(\gamma)$ has appreciable values only within a range $0<\gamma<\gamma_c$ where $\gamma_c \ll k$. Similarly, any individual beam mode has only physical reality if its amplitude distribution function

$$f_{m,n}^{(\pm)}(\gamma) = \left(\frac{\gamma}{\gamma_0}\right)^{m\pm 1} L_{n,m\pm 1}\left(\frac{\gamma^2}{\gamma_0^2}\right) e^{-1/2(\gamma^2/\gamma_0^2)} \tag{8}$$

decays sufficiently rapidly in $\gamma$. Otherwise, the field associated with the mode would not be consistent with Maxwell's equations. Thus, there is an upper bound for $\gamma_0/k$, which depends on the order $n$, $m$ of the Laguerre polynomial.

Each beam mode, as defined in (5), has a plane phase surface at $z=0$. The phase distribution in a plane $z\neq 0$ is given by the factor

$$\exp\left\{j\left[(2n+m\pm 1+1)\operatorname{arctg} v - \frac{1}{2}\frac{u^2 v}{1+v^2}\right]\right\} = \exp\{j[\bar{\psi}_{m,n}^{(\pm)}(z) + \hat{\psi}(\rho, z)]\}, \tag{9}$$

where

$$\bar{\psi}_{m,n}^{(\pm)}(z) = (2n+m\pm 1+1) \operatorname{arc\,tg}\left(\frac{\gamma_0^2}{k} z\right) \tag{9a}$$

is independent of $\rho$ and

$$\hat{\psi}(\rho, z) = -\frac{1}{2} \frac{\frac{\gamma_0^2}{k} z}{1+\left(\frac{\gamma_0^2}{k} z\right)^2} \gamma_0^2 \rho^2 \tag{9b}$$

is independent of the order $n$, $m$ of the Laguerre polynomials. If a phase transformation is performed at the plane $z=+z_0$ by a factor $e^{-i2\hat{\psi}(\rho,z_0)}$ the field distribution at $z=-z_0$ of each beam mode is repeated at $z=+z_0$, disregarding the constant phase shift $2\bar{\psi}_{m,n}^{(\pm)}(z_0)$. Thus, any elementary beam can be restored at intervals $2z_0$ by phase correcting means which effect a radially dependent phase shift

$$\psi(\rho) = -2\hat{\psi}(\rho, z_0) + \psi_0 = +\frac{\frac{\gamma_0^2}{k} z_0}{1+\left(\frac{\gamma_0^2}{k} z_0\right)^2} \gamma_0^2 \rho^2 + \psi_0, \tag{10}$$

where $\psi_0$ is an arbitrary constant.

Although every single beam mode is iterated by the same phase transformation, (10), wave beams composed of several modes are not iterated because the independent phase shift (9a) is different for every mode. A system of equally spaced identical phase transformers guide such a beam, *i.e.*, prevent the energy from spreading out, but the field distribution is not repeated from interval to interval. The conditions are similar to those in an ordinary waveguide where a field consisting of several modes is guided with continuously varying cross-sectional field distribution.

The concentration of energy in a beam mode can be characterized by the fraction $\eta_{m,n}^{(\pm)}$ of the total energy which passes through a plane $z=$ constant within a circle of radius $\rho=\rho_\eta$. If a certain $\eta_{m,n}^{(\pm)}$ is considered, $\rho_\eta$ varies with $z$. The $\rho_\eta(z)$ which characterizes the envelope of

constant energy flux can be derived from (5). Substituting $x=u^2/1+v^2$, one obtains

$$\eta_{m,n}^{(\pm)} = \frac{n!}{(n+m\pm 1)!}\int_0^{x_0} x^{m\pm 1}\{L_{n,m\pm 1}(x)\}^2 e^{-x}dx \quad (11)$$

with

$$x_0 = \frac{\gamma_0^2\rho_\eta^2}{1+\left(\frac{\gamma_0^2}{k}z\right)^2}.$$

This equation shows that the dependency of $\eta_{m,n}^{(\pm)}$ on $\rho_\eta$ and $z$ is only contained in $x_0$. The relation between $\rho_\eta$ and $z$ for $\eta_{m,n}^{(\pm)}=$constant, *i. e.*, $x_0=$constant, can be normalized to

$$\left(\frac{\rho_\eta}{R_0}\right)^2 = \frac{1}{v_0}+v_0\left(\frac{z}{z_0}\right)^2 \quad (12)$$

with

$$v_0 = \frac{\gamma_0^2}{k}z_0, \qquad R_0 = \sqrt{x_0\frac{z_0}{k}}.$$

$\rho_\eta/R_0$ as a function of $z/z_0$ with $v_0$ as parameter is plotted in Fig. 2. If the distance $2z_0$ between two successive phase transformations and the parameter $\gamma_0$ characterizing the set of beam modes are given, all the curves $\rho_\eta(z)$ are similar and are represented by one of the curves of Fig. 2. In other words, assuming the same $z_0$, each curve describes the energy flux distribution for the entire mode system with the parameter $\gamma_0$.

The question can be asked: What value of $\gamma_0$ yields the highest energy concentration, *i.e.*, the smallest "field diameter" at the locations of the phase transformers, if their mutual spacing is given? The desired $\gamma_0$ value is obtained with (12) by introducing the requirement that $\rho_\eta/R_0$, as a function of $v_0$, is a minimum for $z=z_0$. The requirement is satisfied, if $v_0=(\gamma_0^2/k)z_0=1$.

This condition for maximum energy concentration is independent of $n$ and $m$ and, therefore, is the same for each beam mode. The corresponding phase transformation is

$$\psi(\rho) = +\frac{k}{2z_0}\rho^2+\psi_0. \quad (13)$$

The case of maximum energy concentration is illustrated in Fig. 3 which shows various curves $\eta_{1,0}^{(-)}=$constant for the beam mode $m=1$, $n=0$, $(-)$. As a numerical example, we consider a beam at the frequency of 35,000 Mc, whose phase is adjusted at intervals of 10 m ($z_0=5$m). The diameter of the beam for 90 per cent energy flux varies from 25 cm ($z=0$) to 35.4 cm ($z=z_0$), for 99 per cent from 35.6 cm to 50.4 cm, and for 99.9 per cent from 43.6 cm to 61.6 cm.

## Justification of the Simplification Underlying the Theory

The simplified field equations which led to the beam mode system had been obtained with the assumption that the amplitude spectrum $f_m^{(\pm)}(\gamma)$ has considerable values only for $\gamma\ll k$. Since the amplitude function for a beam mode as introduced in the preceding section is

$$f_{m,n}^{(\pm)}(\gamma) = \left(\frac{\gamma}{\gamma_0}\right)^{m\pm 1}L_{n,m\pm 1}\left(\frac{\gamma^2}{\gamma_0^2}\right)e^{-1/2(\gamma^2/\gamma_0^2)}, \quad (14)$$

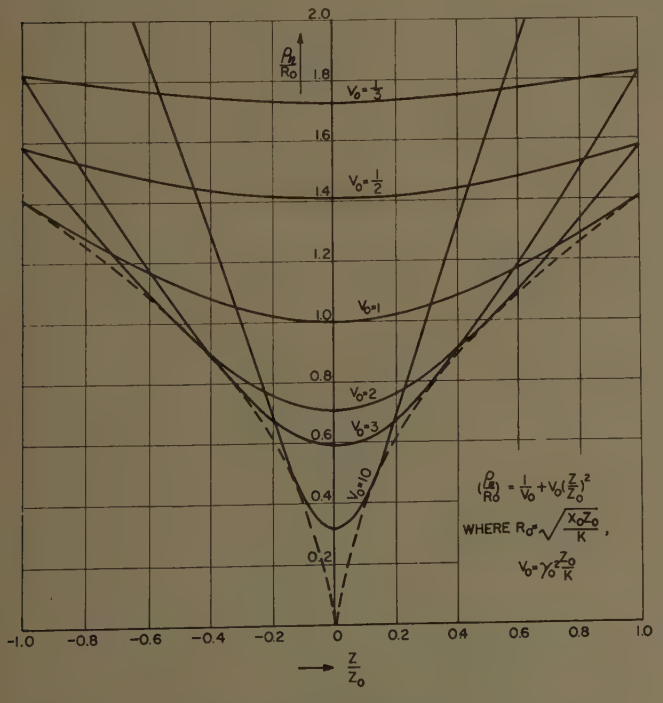



Fig. 2—Envelope curves of constant energy flux in normalized coordinates.

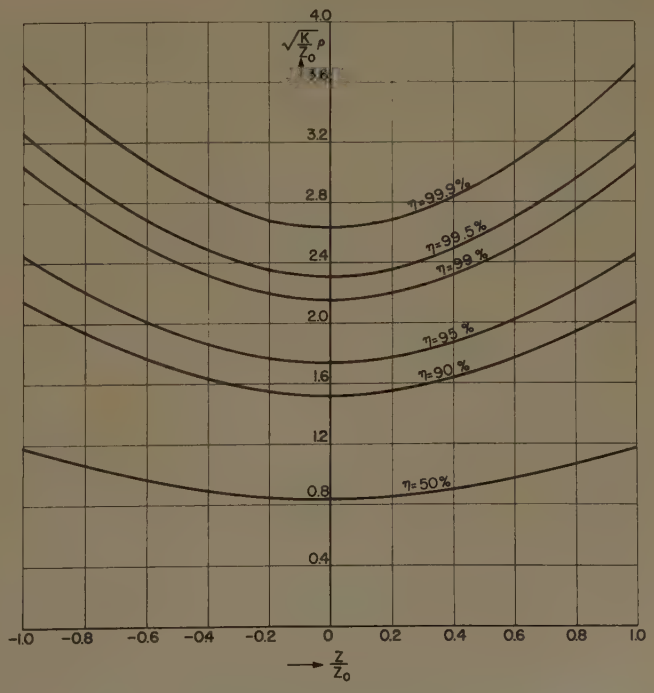



Fig. 3—Envelope curves of constant energy flux for the beam mode $n=0$, $m=1$, $(-)$ in the case of maximum energy concentration at $z=\pm z_0$.

where $L_{n,m\pm1}(\gamma^2/\gamma_0^2)$ is a polynomial in $\gamma^2/\gamma_0^2$, this assumption requires that $\gamma_0$ is sufficiently small. The field distribution of the mode in the plane $z=0$, where the field has the highest concentration, is, according to (5),

$$E_{\rho m,n}^{(\pm)} = \pm E_{\phi m,n}^{(\pm)} = (-1)^n k\gamma_0^2(\gamma_0\rho)^{m\pm1}L_{n,m\pm1}(\gamma_0^2\rho^2)e^{-1/2(\gamma_0^2\rho^2)}. \quad (15)$$

This equation shows that the condition of small $\gamma_0$ implies large beam diameters.

In order to get an estimate as to the meaning of "large beam diameters," consider a beam mode with the parameters $m=1$, $n=0$, $(-)$, where

$$f_{1,0}^{(-)}(\gamma) = e^{-1/2(\gamma^2/\gamma_0^2)} \qquad E_{\rho 1,0}^{(-)} = -E_{\phi 1,0}^{(-)} = k\gamma_0^2 e^{-1/2(\gamma_0^2\rho^2)} \quad \text{at } z=0. \quad (16)$$

If 99 per cent of the energy pass the $z=0$ plane within a circle of radius $\rho=5\lambda$, the corresponding value of $\gamma_0/k$, as obtained with (11), is $\gamma_0/k=6.8\times10^{-2}$. Thus,

$$f_{1,0}^{(-)}(\gamma) = e^{-110(\gamma^2/k^2)}. \quad (17)$$

At $\gamma/k=0.25$, $f_{1,0}{}^{(-)}(\gamma)$ is already as low as $10^{-3}$. If $f_{1,0}{}^{(-)}(\gamma)$ is assumed to be zero for $\gamma/k>0.25$, the relative error of the energy integrals

$$\epsilon\int_F |E_{0,1}^{(-)}|^2 dF = \mu\int_F |H_{0,1}^{(-)}|^2 dF$$

is 0.23 per cent. The phase error introduced into the phase term when $h$ is approximated by $k-\gamma^2/2k$ is below 0.12 radian at $z=z_0$. These errors decrease rapidly if the beam radius is increased. Higher modes require larger radii for the same magnitude in errors.

### Effect of Cross-Sectional Limitations

In the preceding discussion, the assumption had been made that the phase transformation is performed over the entire plane $z=z_0$ and corresponding planes at intervals of $2z_0$.

Actually, the phase transformation can only be achieved over a finite area. A practical solution for the phase transformation is the use of dielectric lenses, as will be explained in another paper.[5]

Although the field of the beam modes, particularly the lowest one, decreases rapidly in the radial direction, some distortion will be introduced if the phase is corrected only within a finite area. The energy outside this area can be considered lost, since the uncorrected field diverges and is radiated into space. The beam mode system which has been derived with the assumption of unlimited phase correction will be modified in two respects: 1) The field functions will depend on the dimensions of the phase correcting devices, and 2) there will be an attenuation factor which is caused by the loss of energy bypassing the phase transformers. This attenuation factor does not include any absorption or reflection losses which may occur in actual phase correcting structures.

In the following, the assumption is made that the phase correction at the planes $z=(\pm2i+1)z_0$, where $i=0, 1, 2, \cdots$, is performed within a circle of radius $R$, while the area outside this circle is covered with an ideal absorber which prevents the field from being reflected or transmitted into the succeeding space section. The modes of this system are characterized by the condition that the field distribution in successive phase correction planes $z=(\pm2i-1)z_0$ and $z=(\pm2i+1)z_0$ is the same, except for a complex constant whose absolute value determines the loss caused by the finite area of phase correction.

The field at $z=-z_0$, after being subjected to the phase transformation, is according to (2)

$$E_{\rho,m}^{(\pm)} = \pm E_{\phi,m}^{(\pm)} = E_m^{(\pm)}(\rho, -z_0) = ke^{+jkz_0}\int_0^\infty f_m^{(\pm)}(\gamma)J_{m\pm1}(\gamma\rho)e^{-j(\gamma^2/2k)z_0}\gamma d\gamma, \quad (18)$$

where $f_m{}^{(\pm)}(\gamma)$ may be complex. $E_m{}^{(\pm)}(\rho, -z_0)$ can be considered as the Hankel transform of $f_m{}^{(\pm)}(\gamma)e^{j-(\gamma^2/2k)z_0}$ of the order $m\pm1$.[6] Therefore, if $E_m{}^{(\pm)}(\rho, -z_0)$ is given $f_m{}^{(\pm)}(\gamma)$ is obtained by the inverse transformation

$$f_m^{(\pm)}(\gamma) = \frac{1}{k}e^{-jkz_0(1-\gamma^2/2k)}\int_0^R E_m^{(\pm)}(\rho, -z_0)J_{m\pm1}(\gamma\rho)\rho d\rho. \quad (19)$$

The integration is only extended from zero to $R$ instead of infinity, because $E_m{}^{(\pm)}(\rho, -z_0)$ is limited to this range by the absorbing screen.

The field at $z=+z_0$ before being subjected to the phase correction is obtained by inserting $f_m{}^{(\pm)}(\gamma)$ of (19) in (2)

$$E_m^{(\pm)}(\rho, +z_0) = e^{-j2kz_0}\int_{\gamma=0}^\infty\int_{\eta=0}^R E_m^{(\pm)}(\eta, -z_0)J_{m\pm1}(\gamma\eta)\cdot J_{m\pm1}(\gamma\rho)e^{j(\gamma^2/k)z_0}\eta d\eta\gamma d\gamma. \quad (20)$$

The integration over $\gamma$ can be performed

$$\int_0^\infty J_{m\pm1}(\gamma\rho)\cdot J_{m\pm1}(\gamma\eta)e^{j(\gamma^2/k)z_0}\gamma d\gamma = j\frac{k}{2z_0}J_{m\pm1}\left(\frac{k}{2z_0}\eta\rho\right)e^{-j(k/4z_0)(\eta^2+\eta^2)}. \quad (21)$$

Thus,

$$E_m^{(\pm)}(\rho, +z_0) = j\frac{k}{2z_0}e^{-j2kz_0}\int_0^R E_m^{(\pm)}(\eta, -z_0)\mathfrak{F}_{m\pm1}\left(\frac{k}{2z_0}\eta\rho\right)\cdot e^{-j(k/4z_0)(\eta^2+\rho^2)}\eta d\eta. \quad (22)$$

[5] J. R. Christian and G. Goubau, "Experimental studies on a beam waveguide for millimeter waves," this issue, pp. 256–263.

[6] W. Magnus and F. Oberhettinger, *op. cit.*, pp. 136–137.

This field, after being subjected to the phase correction, becomes zero in the range $\rho > R$, which is covered by the absorbing screen. The field is somewhat modified within the range $0 < \rho < R$ because of the boundary conditions. However, since we consider beams of large diameters, the modification is limited to a narrow zone along the rim $\rho = R$ and may be neglected. Hence, the field behind the phase transformer is obtained by applying the phase transformation to the center portion $0 < \rho < R$ of the incident field (22).

Obviously the effects of the limitation of the beam will be minimized if the phase correction is adapted to maximum field concentration at $z = z_0$, as given in (13). Then, the field at $z = z_0$, after the phase has been transformed, is

$$E_m^{(\pm)}(\rho, +z_0)e^{j((k/2z_0)\rho^2+\psi_0)} \qquad 0 < \rho < R. \tag{23}$$

This field shall be the same as the field $E_m^{(\pm)}(\rho, -z_0)$, except for a constant complex factor $p$. Thus, we can write

$$E_m^{(\pm)}(\rho, +z_0)e^{j((k/2z_0)\rho^2+\psi_0)} = \rho E_m^{(\pm)}(p, -z_0) \quad \text{for } 0 < \rho < R. \tag{24}$$

Inserted into (20) we obtain

$$pE_m^{(\pm)}(\rho, -z_0)e^{-j(k/4z_0)\rho^2} = j\frac{k}{2z_0}e^{-j(2kz_0+\psi_0)}\int_0^R E_m^{(\pm)}(\eta, -z_0)e^{-j(k/4z_0)\eta^2} \cdot J_{m\pm1}\left(\frac{k}{2z_0}\eta\rho\right)\eta d\eta. \tag{25}$$

We introduce the following abbreviations:

$$E_m^{(\pm)}(\rho, -z_0)e^{-j(k/4z_0)\rho^2} = F_m^{(\pm)}(x), \quad j\frac{1}{p}e^{-j(2kz_0+\psi_0)} = q$$

$$x = \sqrt{\frac{k}{2z_0}}\rho, \quad \xi = \sqrt{\frac{k}{2z_0}}\eta, \quad a = \sqrt{\frac{k}{2z_0}}R. \tag{26}$$

Then (25) transforms into the following normalized homogeneous integral equation of the second kind[7]

$$F_m^{(\pm)}(x) = q\int_0^a F_m^{(\pm)}(\xi)J_{m\pm1}(x\xi)\xi d\xi, \tag{27}$$

whose eigenfunctions and eigenvalues are real and only depend on the quantity $a$.[8]

If the phase corrections at $z = (\pm 2i+1)z_0$ are thought to be performed in two equal steps, then, $F_m^{(\pm)}(x)$ represents the field distribution in between these two steps. Since $F_m^{(\pm)}(x)$ is real, the field in the middle planes of the phase transformers has uniform phase.

The eigenfunctions $F_m^{(+)}(x)$ and $F_m^{(-)}(x)$ form each a complete orthogonal system. Together they determine an orthogonal system for the transverse field vector. They characterize the modes of the system with cross-sectional beam limitation. The eigenvalues $q$ are greater than one. $1/|q| = |p|$ represents the attenuation of the field between two successive phase corrections. Any beam of arbitrary cross-sectional field distribution can be considered as the superposition of these modes. If such a beam is passed through a large number of phase transformers, the "higher modes" which correspond to the higher eigenvalues are gradually damped out until only the lowest mode remains.

If $a$ approaches infinity, the modes transform into the unlimited beam modes. The eigenvalues become $\pm 1$, since there is no beam limitation and, therefore, no diffraction loss.

The following discussion is restricted to the mode group having the parameters $m = 1$, $(-)$. These modes are polarized in one direction. They are the only modes whose transverse field is not zero at the axis.

In order to solve the integral (27) one can proceed in the following manner:

The kernel $J_0(x\xi)$ (for $m = 1$, $(-)$) can be expanded into a series of the mode functions of the unrestricted system

$$J_0(x\xi) = 2\sum_{n=0}^{\infty}(-1)^n L_n(x^2)\cdot L_n(\xi^2)e^{-1/2(x^2+\xi^2)}, \tag{28}$$

where

$$L_n = L_{n,0}.$$

The series converges for any $x$ and $\xi$, except zero and infinity. At $x = \xi = 0$, the series alternates between zero and two.

An approximate solution of the problem can be obtained if $J_0(x\xi)$ is replaced by the first $N$ terms of the series. The resulting approximated integral equation can be solved rigorously. Expanding $F_1^{(-)}(x) = F(x)$ also into the mode functions

$$F(x) = \sum_{n=0}^{N} a_n L_n(x^2)e^{-(1/2)x^2}, \tag{29}$$

the approximated integral equation yields

$$\sum_{n=0}^{N} a_n L_n(x^2)e^{-(1/2)x} = q\sum_{r=0}^{N}\sum_{n=0}^{N}(-1)^n b_{nr}a_r L_r(x^2)e^{-(1/2)x^2}, \tag{30}$$

where

$$b_{nr} = 2\int_0^a L_n(x^2)L_r(x^2)e^{-x^2}x dx,$$

[7] This equation can also be derived with Fresnel-Kircheoff's diffraction theory. This approach has been used by G. Goubau and R. Christian, "A New Waveguide for Millimeter Waves," presented at the URSI-IRE Fall Meeting, San Diego, Calif.; October 19, 1959.

[8] R. Courant and D. Hilbert, "Methoden die Mathematik Physik," Springer-Verlag, Berlin, Ger., vol. 1, pp. 96–138; 1937.

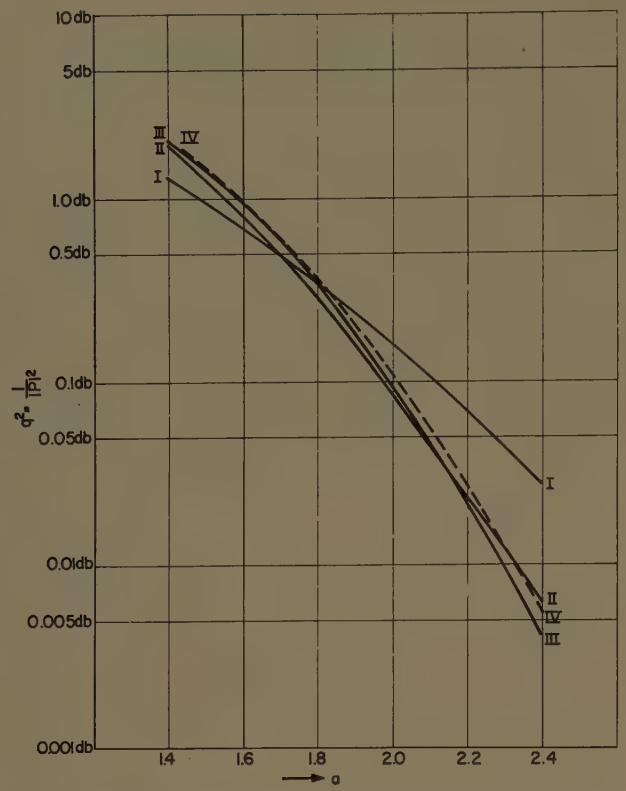



Fig. 4—The first eigenvalue of integral (23). Curves I, II and III are obtained by expanding the kernel $J_0(x\xi)$ into the system of mode functions using 1, 3 and 5 terms, respectively. Curve IV has been computed with an IBM computer applying power series development.

(30) resolves into a system of linear homogeneous equations for the coefficients $a_n$, since the equation must be satisfied for any $x$

$$\sum_{n=0}^{N} (b_{nr} - (-1)^n q\delta_{nr})a_n = 0 \qquad 0 \leq r \leq N,$$

$$\delta_{nr} = \begin{matrix} 1 & r = n \\ 0 & r \neq n. \end{matrix} \qquad (31)$$

The determinant of this system must be zero. This condition leads to an algebraic equation of $N$th degree. The roots are the desired approximations for the eigenvalues. The corresponding coefficients $a_n$ determine the approximated eigenfunctions.

Fig. 4 shows curves for the first eigenvalue as a function of $a=\sqrt{(k/2z_0)}R$ calculated in this manner for $N=1$, 3 and 5 (curves I to III). Curve IV has been obtained with an IBM computer using for $J_0(x\xi)$ a power series expansion. The ordinate is given in a db scale and refers to $p^2$ which determines the diffraction loss caused by the finite apperture. Fig. 5 is a plot of eigenfunctions for $a=1.8$, 2.0, 2.2 and 2.4, as obtained with the computer.

In order to obtain an estimate for the attenuation of the higher modes, the eigenvalues for $n=2$ and 3 have been calculated with $N=3$. The results are given in Table I. Presumably, these values are not very accurate, since higher modes require larger $N$ for the same degree of accuracy.

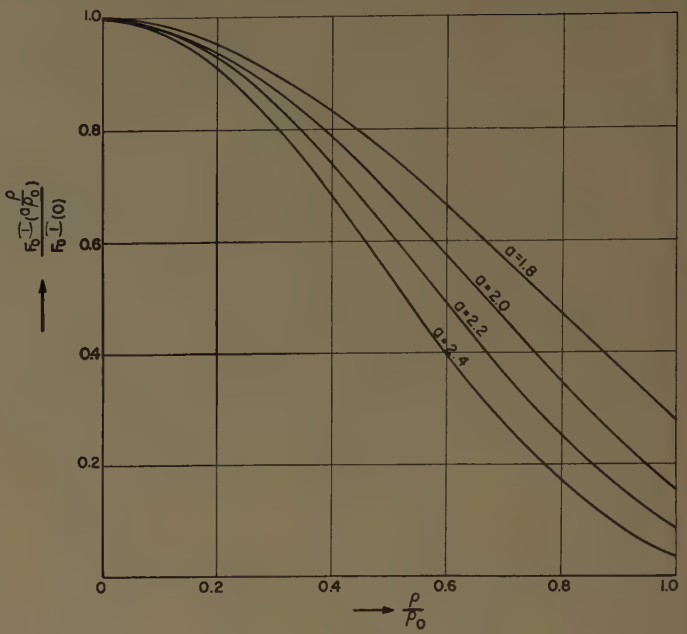



Fig. 5—The first eigenfunction of integral (23) for various values of $a$.

TABLE I

| $a$ | $p(n=1)$ | $p(n=2)$ | $p(n=3)$ |
|---|---|---|---|
| 1.4 | 0.797 | −0.040 | 0.102 |
| 1.6 | 0.911 | −0.142 | 0.113 |
| 1.8 | 0.968 | −0.324 | 0.158 |
| 2.0 | 0.990 | −0.538 | 0.236 |
| 2.2 | 0.997 | −0.727 | 9.348 |
| 2.4 | 0.999 | −0.859 | 0.487 |

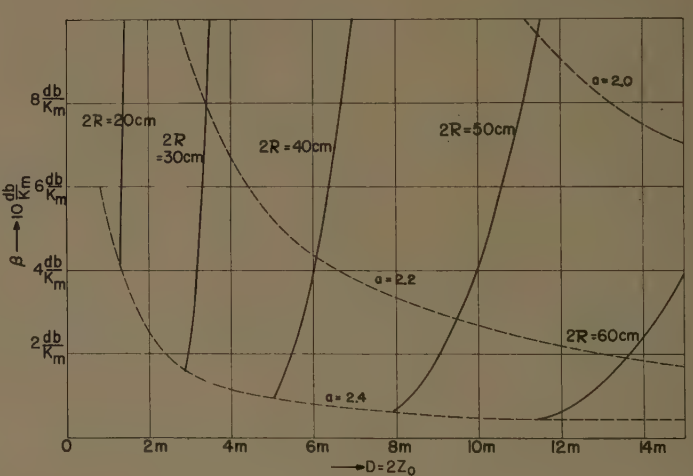



Fig. 6—Calculated diffraction loss per km of a beam waveguide at 35,000 Mc for various beam diameters $2R$ as a function of the transformer spacing $D=2z_0$.

The diffraction loss of a beam at a frequency of 35,000 Mc, guided by an array of phase transformers over a distance of 1 km is shown in Fig. 6. Each curve refers to a certain diameter of the beam measured at the locations of phase transformation. The loss is very small as long as $a \geqq 2.4$ and increases rapidly with decreasing $a$, *i.e.*, with increasing $z_0$. The spacing increases for the

same amount of loss more than with the square of the beam diameter.

## Appendix I

### *Ideal Phase Transformer*

Assume at $z=z_0$ a thin layer (thickness $\delta$) of dielectric material whose $\epsilon$ and $\mu$ are functions of $\rho$ and $\phi$, but not of $z$. A wave beam of the kind described by (12) shall pass through the layer without producing a reflected field and, thereby, change its phase distribution in a predescribed manner.

The transverse field vectors $\boldsymbol{E}_t$ and $\boldsymbol{H}_t$ at both sides of the layer are

$$\boldsymbol{E}_t = \boldsymbol{E}(\rho,\phi) \qquad \sqrt{\frac{\mu}{\epsilon}}\boldsymbol{H}_t = e_z \times \boldsymbol{E}(\rho,\phi) \quad \text{for } z = z_0,$$

$$\boldsymbol{E}_t = \boldsymbol{E}(\rho,\phi)e^{j\alpha(\rho,\phi)}, \quad \sqrt{\frac{\mu}{\epsilon}}\boldsymbol{H}_t = e_z \times \boldsymbol{E}(\rho,\phi)e^{j\alpha(\rho,\phi)} \quad \text{for } z = z_0+\delta. \tag{32}$$

$\alpha(\rho,\phi)$ is the desired phase shift which is a function of $\rho$ and $\phi$. $e_z$ is the unit vector in the $z$-direction.

Since the transverse field components are continuous at the interface of any two media, $\boldsymbol{E}_t$ and $\boldsymbol{H}_t$ are the same inside the layer. The longitudinal field components of the incident field are reduced by the factors $\epsilon/\epsilon_1$ and $\mu/\mu_1$, respectively, after the field penetrates the interface. If $\epsilon_1$ and $\mu_1$ are sufficiently large, the longitudinal field components within the layer are negligible. Thus, in spite of the variation of the field with $\rho$ and $\phi$, the field inside the layer can be treated as an ordinary plane wave. The condition for no reflection at the interfaces is $\epsilon/\mu=\epsilon_1/\mu_1$. The required phase shift inside the layer is $\alpha(\rho,\phi)=-\omega\sqrt{\mu_1\epsilon_1}\delta+2\pi n$. The distribution functions of $\epsilon_1$ and $\mu_1$ which are required to effect this phase shift are, therefore, given by

$$\epsilon_1(\rho,\phi)\delta = \frac{1}{\omega}\sqrt{\frac{\epsilon}{\mu}}\,(2\pi n - \alpha(\rho,\phi)), \tag{33}$$

$$\mu_1(\rho,\phi)\delta = \frac{1}{\omega}\sqrt{\frac{\mu}{\epsilon}}\,(2\pi n - \alpha(\rho,\phi)). \tag{34}$$

Since the phase shift depends on the products $\epsilon_1\cdot\delta$ and $\mu_1\cdot\delta$, the thickness $\delta$ can be made arbitrarily small if $\epsilon_1$ and $\mu_1$ are sufficiently large. Instead of varying $\epsilon_1$ and $\mu_1$, the thickness $\delta$ may depend on $\rho$ and $\phi$. The integer $n$ may have different values for different ranges of $\rho$.

## Appendix II

### *Proof of the Orthogonality of Beam Modes*

We first normalize the beam modes by setting

$$A_{m,n}^{(\pm)} = \gamma_0\left\{\frac{n!}{\pi(n+m\pm 1)!}\right\}^{1/2}. \tag{35}$$

First (6b) shall be proven. The integrals in these equations are extended over a plane $z$-constant. Thus,

$$\int_F \boldsymbol{E}_{m,n}^{(\pm)}\cdot\boldsymbol{E}_{m',n'}^{(\pm)*}dF = \int_{\rho=0}^{\infty}\int_{\phi=0}^{2\pi}\left\{E_{\rho m,n}^{(\pm)}E_{\rho m',n'}^{(\pm)*} + E_{\phi m,n}^{(\pm)}\cdot E_{\phi m',n'}^{(\pm)*}\right\}d\phi\rho d\rho$$

$$\int_F \boldsymbol{E}_{m,n}^{(\pm)}\cdot\boldsymbol{E}_{m',n'}^{(\mp)*}dF = \int_{\rho=0}^{\infty}\int_{\phi=0}^{2\pi}\left\{E_{\rho m,n}^{(\pm)}E_{\rho m',n'}^{(\mp)*} + E_{\phi m,n}^{(\pm)}\cdot E_{\phi m',n'}^{(\mp)*}\right\}d\phi\rho d\rho.$$

$E_{\rho m,n}^{(\pm)}$ and $E_{\phi m,n}^{(\mp)}$ are replaced by the right hand side of (5a), substituting

$$x = \frac{u^2}{1+v^2} = \frac{\gamma_0^2\rho^2}{1+\left(\frac{\gamma_0^2}{k}z\right)^2}$$

we obtain with the normalization (35)

$$\int_f \boldsymbol{E}_{m,n}^{(\pm)}\cdot\boldsymbol{E}_{m',n'}^{(\pm)*}dF = e^{j\{2(n-n')+(m-m')\}\operatorname{arctg} v}\cdot\left\{\frac{n!}{(n+m\pm1)!}\cdot\frac{n'!}{(n'+m'\pm1)!}\right\}^{1/2}\cdot\int_{x=0}^{\infty}x^{1/2(m+m')\pm1}L_{n,m\pm1}(x)L_{n',m'\pm1}(x)e^{-x}dx\cdot\frac{1}{2\pi}\int_0^{2\pi}\cos\left\{(m-m')\phi+(\alpha_m-\alpha_{m'})\right\}d\phi \tag{36}$$

$$\int \boldsymbol{E}_{m,n}^{(\pm)}\cdot\boldsymbol{E}_{m',n'}^{(\mp)*}dF = e^{j\{2(n+n')+(m+m')\pm2+2\}\operatorname{arctg} v}\cdot\left\{\frac{n!}{(n+m\pm1)!}\cdot\frac{n'!}{(n'+m'\mp1)!}\right\}^{1/2}\cdot\int_{x=0}^{\infty}x^{1/2(m+m')}L_{n,m\pm1}(x)L_{n',m'\mp1}(x)e^{-x}dx\cdot\frac{1}{2\pi}\int_0^{2\pi}\cos\left\{(m+m')\phi+(\alpha+\alpha_{m'})\right\}d\phi. \tag{37}$$

Eq. (36) applies to all values of $m$ and $m'$. Eq. (37) does not exist for $m=m'=0$, since it applies to modes with different superscripts $(+)$ and $(-)$, while in the case $m=0$, there are only modes with one superscript which has been chosen to be $(+)$.

The integrations with regard to $\phi$ in (36) and (37) are readily performed

$$\frac{1}{2\pi}\int_0^{2\pi} \cos\{(m-m')\phi + (\alpha_m - \alpha_{m'})\}d\phi = \delta_{m,m'} \quad (38)$$

$$\frac{1}{2\pi}\int_0^{2\pi} \cos\{(m+m')\phi + (\alpha_m + \alpha_{m'})\}d\phi = 0. \quad (39)$$

The orthogonality relations of the Laguerre polynomials yield

$$\int_0^{\infty} x^k L_{n,k}(x) L_{n',k}(x) e^{-x} dx = \frac{(n+k)!}{n!}\delta_{n,n'}. \quad (40)$$

Eq. (39) inserted into (37), and (38) and (40) inserted into (36) prove the orthogonality relations (6b).

The proof of (6a) is now readily obtained, since with (5b) the relations

$$\{\boldsymbol{E}^{(\pm)}_{m,n} \times \boldsymbol{H}^{(\pm)^*}_{m',n'}\}e_z = \sqrt{\frac{\mu}{\epsilon}}\,\boldsymbol{E}^{(\pm)}_{m,n}\cdot\boldsymbol{E}^{(\pm)^*}_{m',n'}$$

$$\{\boldsymbol{E}^{(\pm)}_{m,n} \times \boldsymbol{H}^{(\mp)^*}_{m',n'}\}e_z = \sqrt{\frac{\mu}{\epsilon}}\,\boldsymbol{E}^{(\pm)}_{m,n}\cdot\boldsymbol{E}^{(\mp)^*}_{m',n'} \quad (41)$$

can be easily verified. Thus, (6a) is mathematically identical with (6b).

ACKNOWLEDGMENT

Thanks are due to the Mathematical Division of the Institute of Exploratory Research, USASRDL, Fort Monmouth, N. J., particularly to Dr. I. Epstein and N. Gordon, for the numerical evaluation of the integral equation.

# Experimental Studies on a Beam Waveguide for Millimeter Waves*

J. R. CHRISTIAN†, MEMBER, IRE, AND G. GOUBAU†, FELLOW, IRE

*Summary*—**The beam waveguide utilizes reiterative wave beams which are guided by resetting the cross-sectional phase distributions at periodic intervals. The major purpose of the measurements was to obtain data concerning the inherent losses of the guide, including the diffraction loss, caused by energy by-passing the phase transformers, and the dielectric and reflection losses of the transformers. Resonance measurements with an "open cavity," consisting of a section of a beam waveguide terminated by plane reflector walls, yielded the over-all losses of the phase transformers. In order to isolate the diffraction loss, pulse measurements were made between two reflectors, slightly curved to reset the phase distribution in the reflected beam. These measurements also gave information on the build-up of the reiterative wave beam. The experimental data support the theory which predicts that the diffraction losses of reiterative beams can be made extremely small.**

## I. INTRODUCTION

THE waveguide discussed in this paper differs from usual waveguides in that the energy is not transmitted in the form of a wave mode. A wave mode is characterized by a certain propagation constant and a certain field pattern which is everywhere the same along the guide. The field in a beam waveguide consists of a wave beam, that is, a bundle of waves characterized by a spectrum of propagation constants. The field distribution varies along the guide, but is periodically repeated at certain intervals. The guidance is accomplished by passing the beam through "phase transformers," that is, phase shifting devices which reset the cross-sectional phase distribution in the beam at predetermined intervals.

Beams which can be guided in this manner are called reiterative wave beams. The theory of reiterative beams is contained in a preceding paper.[1] There are sets of reiterative beams called beam "modes" which satisfy the same orthogonality relation as the modes in an ordinary waveguide. Beam modes of the same set require the same phase transformation for their guidance. The attenuation differs for the various modes. The lowest mode whose radial intensity distribution follows essentially an $e^{-\alpha r^2}$ curve, is of particular interest since it has the lowest attenuation for a given beam diameter; it is this mode which is utilized in the beam waveguide.

Fig. 1 shows a schematic drawing of a beam waveguide. It comprises uniformly spaced phase transformers which represent the actual guiding structure and terminations for launching and receiving the desired beam mode. The phase transformers are designed to advance the phase in the outer portion of the beam relative to the center according to the relation

$$\phi = \frac{k\rho^2}{D},$$

* Received by the PGAP, June 16, 1960.
† U. S. Army Signal Res. and Dev. Lab., Fort Monmouth, N. J.

[1] G. Goubau and F. Schwering, "On the guided propagation of electromagnetic wave beams," this issue, pp. 248–256.

where $\phi$ is the phase advance at the radius $\rho$, $D$, the spacing of the phase transformers, and $k=2\pi/\lambda$. The maximum phase shift $\phi_{max}=kR^2/D$ ($R$ is the radius of the phase transformer) is directly related to the diffraction loss $L_D$ of a phase transformer, that is, the loss caused by the cross-sectional limitation of the beam. The relationship between $\phi_{max}$ and $L_D$ is shown in Fig. 2 which is essentially curve IV of Fig. 4 in the above quoted paper.[1] The loss drops below 1/100 db if $\phi_{max}$ approaches $2\pi$. Thus the maximum phase shift for all practical cases is approximately $2\pi$.

Phase shifts which obey a quadratic relation in $\rho$ can be accomplished by use of dielectric lenses. However, it shall be emphasized that the lenses in this application do not function like optical lenses. Geometric optical conditions would require much larger values of $\phi_{max}$, *i.e.*, much higher operational frequencies. The focal length of the lenses which follows from (1) is $D/2$. Under geometric optical conditions each lens would image the preceding lens into the succeeding lens.

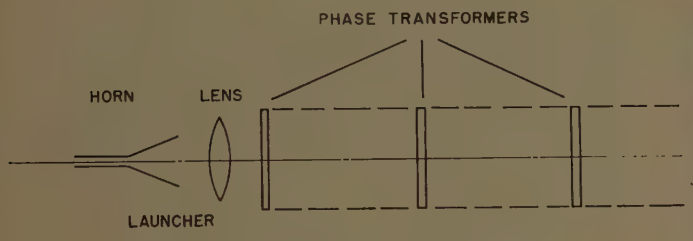



Fig. 1—Schematic of a beam waveguide.

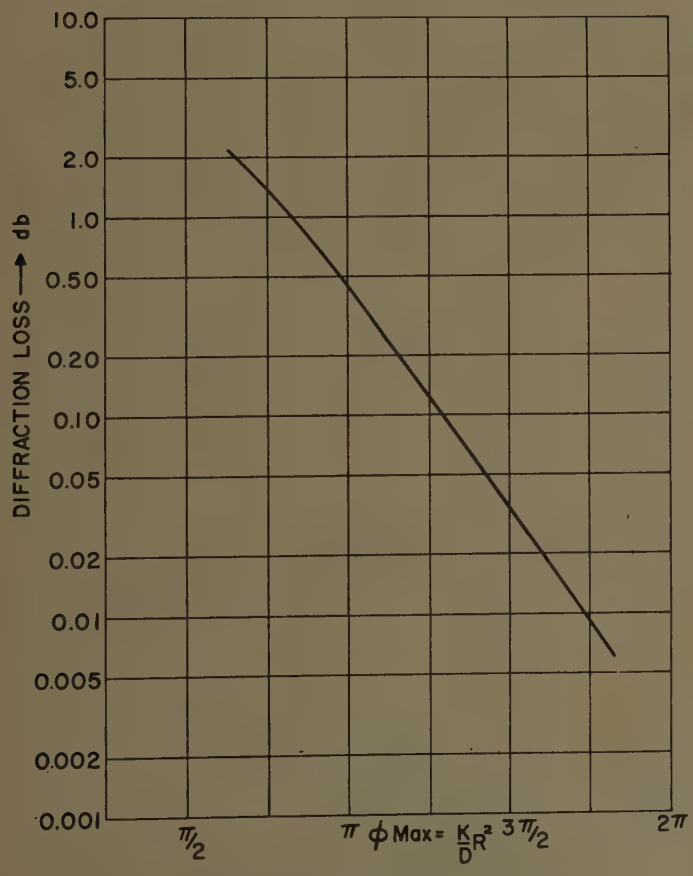



Fig. 2—Calculated diffraction loss as a function of maximum phase shift of a phase transformer.

## II. Modal Beam Waveguide

In order to demonstrate the existence of reiterative wave beams and to obtain some information on the performance of a beam waveguide, a simple model was built for frequencies around 23 kMc. A photograph of this model is shown in Fig. 3. The phase transformers were dielectric lenses set in wooden frames. Their diameter was 20 cm and their mutual spacing approximately one meter. Lauching and receiving were accomplished by means of horns in combination with lenses. The aperture of the horns was circular and had a diameter of 10 cm. Standard K-band rectangular waveguide was used to feed the horns. The wave beam emitted from the horns was substantially linearly polarized.

Two types of phase transformers were used in the model, both employing different methods for minimizing reflection, since reflection causes transmission losses and undesired standing waves.

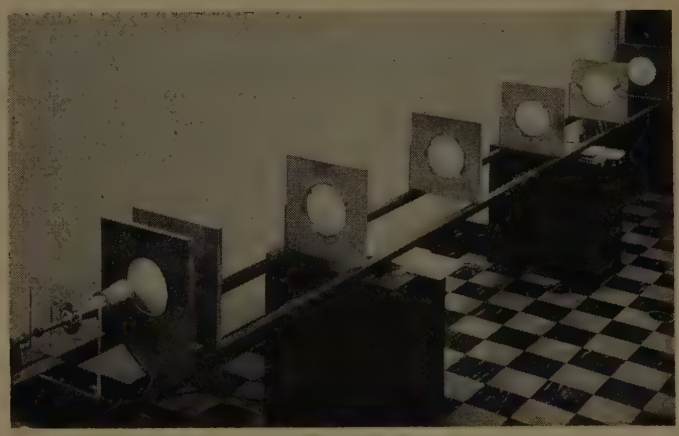


Fig. 3—Photograph of a model beam waveguide.

One type (Fig. 4) employed a single lens of polystyrene foam with a dielectric constant of only 1.2. Because of the low dielectric constant the reflection loss is less than 0.02 db. The dimensions are indicated in Fig. 4.

The other type of phase transformers comprised a pair of identical lenses mounted back to back, as shown in Fig. 5. The design of the lenses was based on the well-known fact that the phase of a plane wave can be shifted without reflection by passing the wave through a pair of appropriately spaced parallel dielectric plates of equal thickness. Each pair of adjacent elements, indicated in Fig. 5 by 1 and 2, has been treated like a pair of parallel dielectric plates effecting the desired phase transformation at the location of these elements. Although such a design procedure is only justified for phase transformers of extremely large diameters, it is apparently applicable to diameters of only 16 wavelengths, since the phase transformer exhibited negligible reflection at their design frequency of 23 kMc.

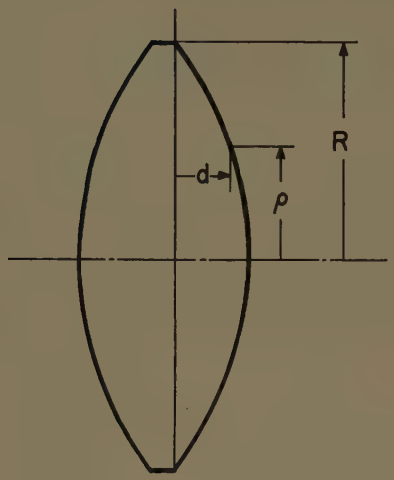



Fig. 4—Foam dielectric phase transformer

$$d = \frac{R^2 - \rho^2}{2D(\sqrt{\epsilon} - 1)}. \quad R = 10\text{ cm},\ D = 1.1\text{ m},\ \epsilon = 1.2.$$

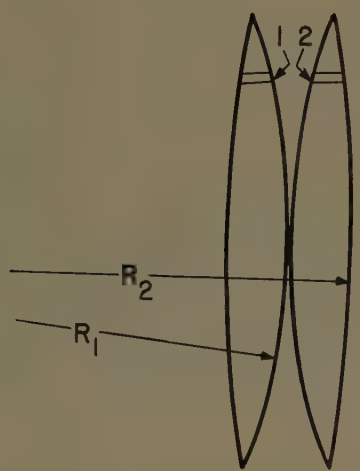



Fig. 5—Double-lens phase transformer.
Radii of curvature, $R_1 = 75$ cm, $R_2 = 216$ cm, $\epsilon = 2.55$.

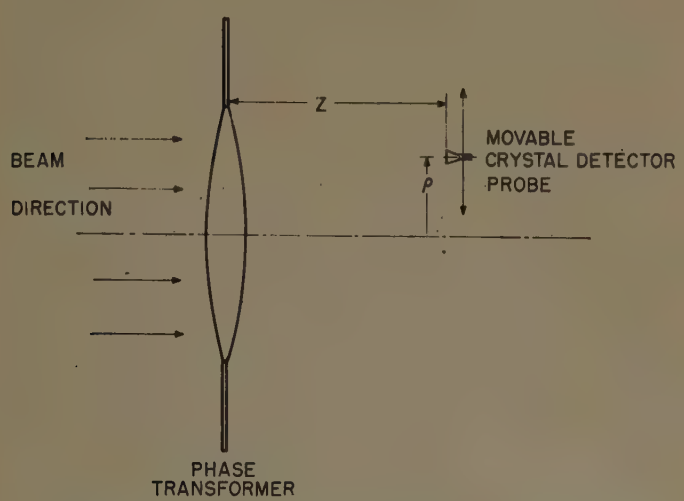



Fig. 6—Schematic of field distribution measurement setup.

A number of field measurements was made to explore the cross-sectional field distribution within the space between the phase transformers. Fig. 6 shows a schematic of the measuring setup. An open-ended flanged waveguide was used as the probe for most measurements. A rectangular horn, 2 cm across its aperture, was used to investigate the lower level field amplitudes (more than 25 db below $E_{max}$). The indicator consisted of a calibrated crystal detector connected to a millivoltmeter.

The field distribution in the first sections of the beam waveguide depended largely on the adjustment of the launcher, that is, on the mutual distances between the horn and the lenses. This adjustment affects the field distribution of the beam developed by the launcher. If this distribution deviates from that of the lowest beam mode which is propagated by the lens structure, higher beam modes are excited simultaneously. In spite of their greater attenuation, the presence of these modes is still noticeable within the first sections of the beam waveguide. It was found that even after the beam had passed eight transformers, the field distribution was still somewhat dependent on the adjustment of the launcher.

In order to obtain representative field measurements of the reiterative wave beam with a small number of available phase transformers (8 and 10 respectively), the launcher was carefully adjusted to minimize the development of higher modes. A criterion for this adjustment was that the field distribution remained substantially unchanged after the beam had passed seven phase transformers.

Fig. 7(a)–7(d) shows measured cross-sectional distributions at four different distances ($D/4$, $D/2$, $3D/4$, $D$) behind the eighth phase transformer. The foam lenses were used in these measurements but the double lens transformers yielded similar results. The abscissa indicates the distance from the beam axis, and the ordinate indicates the ratio $[E(\rho, z)/E(0, D)]^2$ where $E(\rho, z)$ is the field at the distance $\rho$ from the axis and $E(0, D)$ the field in the center of the beam at the distance $D$. The broken line refers to the $E$ plane and the dotted curve to the $H$ plane. According to the theory, both curves should be identical if the beam diameter is sufficiently large. The deviation between the $E$- and $H$-plane curves may have been caused by the fact that the beam diameter of 16 wavelengths was not large enough and/or that the launching horn which radiates differently within the $E$ and $H$ planes causes the excitation of an azimuthally periodic beam mode of also relatively low attenuation.

Fig. 7(d) which refers to the field distribution at the distance $z = D$, that is, the location of a phase transformer, has been measured for a wider cross-sectional range, thus showing the distribution of the field which bypasses the phase transformer. The solid curve is the computed field distribution for $\sqrt{k/D} \cdot R = 2.2$. This curve has been taken from Fig. 5 of Goubau and

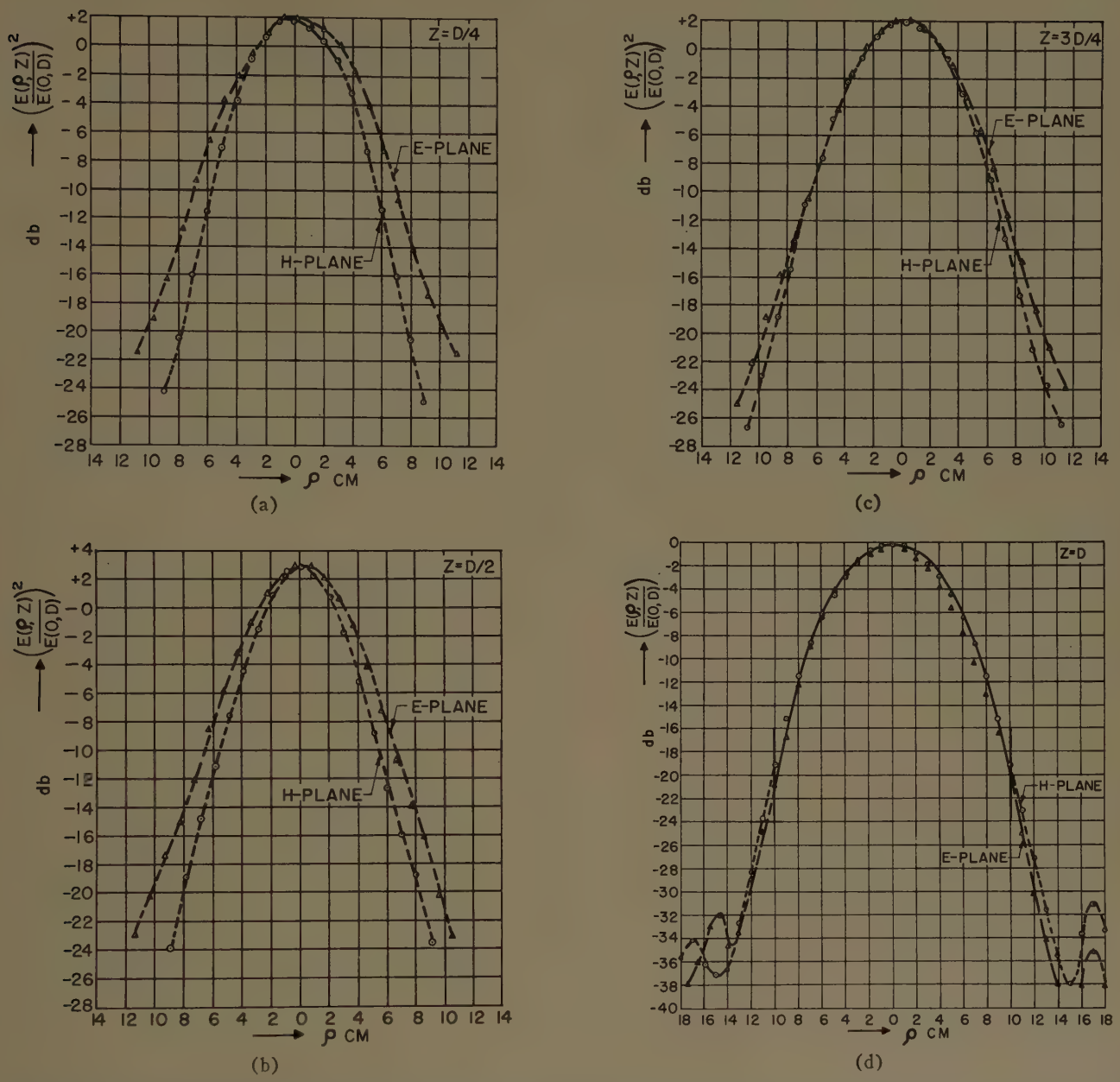



Fig. 7—(a) Cross-sectional field distribution at a distance $z = D/4$ behind the eighth phase transformer. (b)—Cross-sectional field distribution at a distance $z = D/2$ behind the eighth phase transformer. (c)—Cross-sectional field distribution at a distance $z = 3D/4$ behind the eighth phase transformer. (d)—Cross-sectional field distribution at a distance $z = D$ behind the eighth phase transformer.

Schwering.[1] The actual value of $\sqrt{k/D} \cdot R$ for the phase transformers used in the model was somewhat smaller, namely 2.15.

In order to obtain information on the efficiency of the beam waveguide, the over-all transmission loss was measured between the input terminal of the launching horn and the output terminal of the receiving horn. The measured loss was 2.5 db using 8 foam lens transformers and 1.9 db using 10 double lens transformers. In both cases the spacing $D$ was 1.1 meter. The loss remained unchanged when the wooden frames were covered by aluminum. This observation is of particular interest with regard to the theory which ignores specific boundary conditions at the rim of the phase transformers.[1]

The transmission loss consists of two components, the "iteration loss" which is caused by the phase transformers and the "launching loss" caused by the terminations. Since the launching loss is the same in both cases, the iteration loss of the foam lenses is apparently larger than that of the double lens configurations. As will be shown later, the iteration loss of a foam lens was 0.11 db and that of a double lens transformer approximately 0.05 db. Therefore the launching loss was approximately 1.5 db. Since the largest part of the measured transmission loss was launching loss, it was not possible to determine the iteration loss by adding or removing a phase transformer. The adjustment of the terminating horn-lens combinations was quite critical and the variation

of the launching loss was of the same magnitude or even greater than the loss of an individual phase transformer. Less critical was the alignment of the phase transformers, although transverse displacements of the transformers divert the beam from its original path. No noticeable change in loss was produced by a slight tilt of the phase transformers, since a tilt causes only a second-order effect on the phase distribution. It was found that the spacing of the phase transformers could be varied within 10 per cent without affecting the total transmission loss.

### III. Iteration Loss

The iteration loss of a phase transformer consists of the diffraction loss and the reflection and absorption losses. The theoretical diffraction loss can be obtained from Fig. 2 as a function of $(k/D)R^2$. The reflection loss pertains mainly to the phase transformers, which are uncompensated for reflection. The power reflection coefficient $\Gamma$ of one surface of a dielectric lens can be calculated approximately. For values of $\epsilon \geq 1.2$ one obtains

$$\Gamma \approx \left(\frac{\sqrt{\epsilon}-1}{\sqrt{\epsilon}+1}\right)^2 + \left(\frac{1}{4}\,\frac{1}{\epsilon\sqrt{\epsilon}}\,\frac{(4\sqrt{\epsilon}-\epsilon-1)}{(\epsilon-1)^2} \cdot \frac{1}{(kD)^2}\,\frac{1-(1+a^2+a^4/2)e^{-a^2}}{1-e^{-a^2}}\right), \quad (2)$$

where $a=\sqrt{k/D}\cdot R$, and $\epsilon$ is the relative dielectric constant. The first term represents the reflection from a plane dielectric interface; the second term takes into account the curvature of the interface. The reflection loss of a single lens phase transformer, assuring no interaction between the surfaces, is

$$L_R = -2[10\log_{10}(1-\Gamma)]\text{db}. \quad (3)$$

The absorption loss within the dielectric depends not only on the dissipation factor of the dielectric material but also on the field distribution. One obtains for the absorption coefficient the expression

$$\eta \approx \frac{\epsilon_r(a^2-1)}{\sqrt{\epsilon}-1}, \quad (4)$$

where $\epsilon_r$ = dissipation factor. The absorption loss is then given by

$$L_A = -10\log_{10}(1-\eta)\text{db}. \quad (5)$$

The reflection and absorption losses are plotted individually in Fig. 8 as a function of the dielectric constant, assuming a 2.16 value for $a$. Two curves are shown for the absorption loss for values of $\epsilon_r = 0.0001$ and 0.0002. Curves representing the total dielectric loss of the phase transformer, the summation of reflection and absorption loss, are also shown.

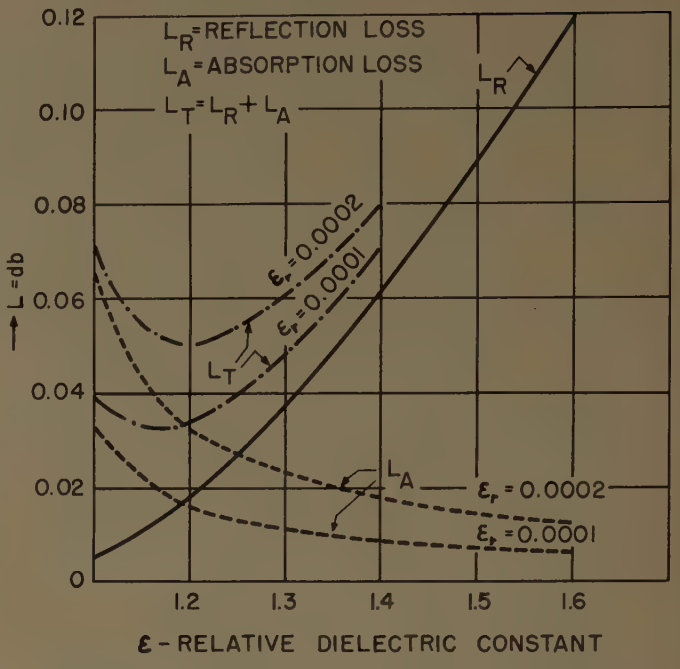



Fig. 8—Phase-transformer dielectric losses as a function of dielectric constant for an $a$ value of 2.16.

In order to eliminate the launching variable and to obtain measured data as to the individual phase-transformer loss, a parallel plate resonator was built using one section of the beam waveguide. This resonator consists of two plane reflector plates, one placed $D/2$ before and the other placed $D/2$ after the phase transformer (Fig. 9). This resonator behaves like a transmission-line section short-circuited at both ends. The resonant frequencies are approximately given by $f=nc/2D$, where $n$ is an integer. By measuring the $Q$ of the resonator, the loss of the phase transformer can be obtained[2] from the expression

$$L = \frac{\pi D}{Q\lambda_0}\,8.68\text{ db}, \quad (6)$$

where $Q=f/\Delta f$, and $\Delta f$ is obtained from the half-power points of the resonance curve. The above expression does not consider a loss in the reflector plates which is indeed negligible.

Since $Q$'s in the order of 20,000 or more were expected, a very stable oscillator was required. An atomicron stabilized Klystron oscillator which had a possible frequency variation of $\pm 3$ Mc was employed to obtain the $Q$ measurement. The coupling to the resonator was performed at the center of each plate through small slots. Measurements were made with the foam phase transformers. The double-lens transformer unfortunately could not be measured in the experimental set up because its reflection compensation occurred at a fre-

[2] See, for instance, E. L. Ginzton, "Microwave Measurements," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 469–470; 1957.

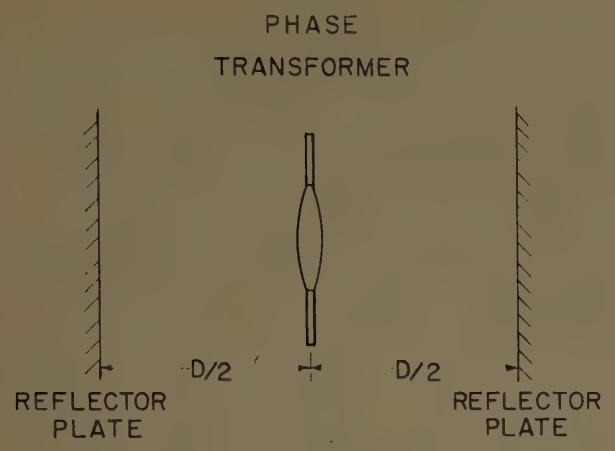



Fig. 9—Schematic of parallel-plate resonator.

quency which was outside the range of the measuring equipment. In order to measure $Q$ as a function of $a$ the radius $R$ of the phase transformer was varied by inserting aluminum plates with circular apertures. The points in Fig. 10 show the loss $L$ calculated from the measured $Q$ value as a function of $a$. Also shown as solid curves are theoretical iteration losses assuming dissipation factors of $\epsilon_r = 0.0002$, 0.0003, 0.0004. The measured values are in good agreement with the curve for $\epsilon_r = 0.0003$.[3]

## IV. Build-Up of Reiterative Wave Beams

The experiments conducted on the model beam waveguide indicated that the wave beam had to pass through several phase transformers before the repetitive beam mode was established. In order to investigate the build-up of the reiterative beams, pulse-type measurements were conducted on a specially constructed beam waveguide resonator (Fig. 11). This resonator differed from the one used in the preceding experiment, in that the reflector plates were, in effect, located within the center planes of two adjacent transformers. The advantage of this arrangement is that the required phase transformation can be achieved by appropriately curving the reflector plates rather than by using dielectric lenses, thus eliminating the dielectric losses. In other words, the line section simply consisted of two curved reflector plates facing each other as shown.

The radii of curvature of the reflector plates coincided with their spacing which was approximately 30 meters. The diameter of each plate was 0.91 meters, corresponding to a value of $\phi_{max} = 1.6\pi$. The reflector plates were machined from lucite sheet stock, the final surfaces being plated with silver. The plates were mounted on steel frames, with a provision made on the back of these frames to allow a small deformation of the reflector plates. In this manner it was possible to adjust the average curvature of the reflectors which varied slightly with temperature.

[3] The manufacturer of the foam material, Emerson and Cuming, Inc., Canton, Mass., quotes an average dissipation factor of 0.0002 for a frequency range of $10^9$ to $3\times10^{10}$ cps.

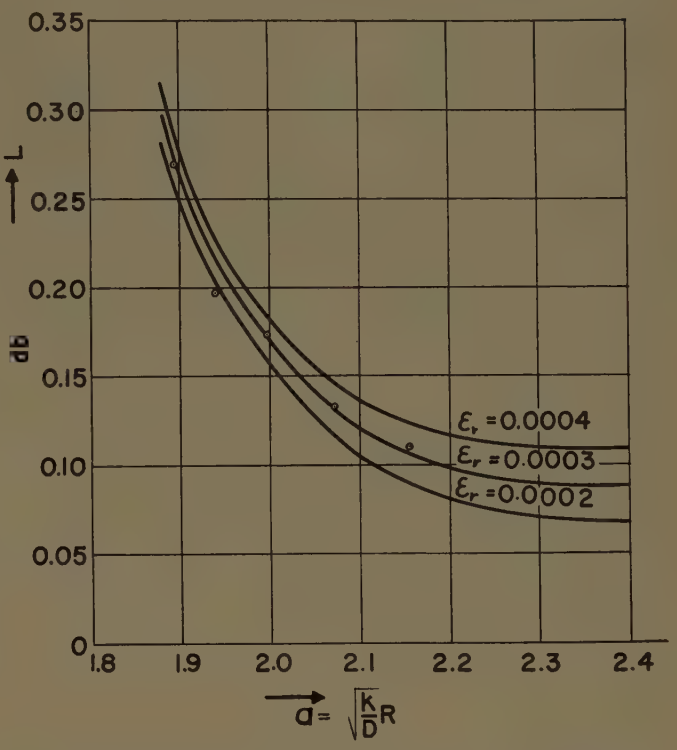



Fig. 10—Measured and calculated iteration loss of a foam dielectric phase transformer as a function of $a = \sqrt{(k/D)}R$.

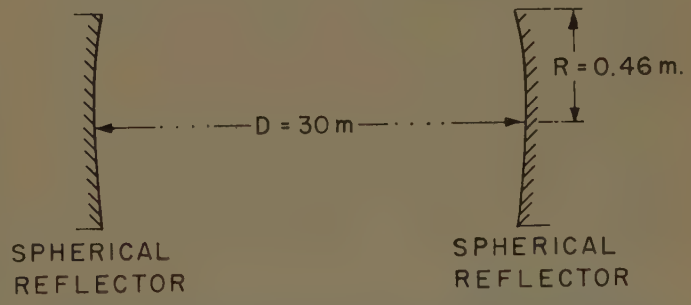



Fig. 11—Experimental setup for determining diffraction loss.

A millimeter wave communication set was used as a signal generator, supplying 35 kMc pulses of 70 mμsec. duration. The pulses were introduced into the cavity through a small coupling hole in the center of one of the reflector plates. The coupling hole, a square aperture of one cm on each side, was fed from standard rectangular waveguide. The same coupling scheme was used on the receiving side. The receiving system employed a crystal detector followed by a series of wide-band chain amplifiers, the output of which was displayed on an oscilloscope. In order to view any portion of the series of received pulses, a variable delay network was inserted between the oscilloscope and the transmitter trigger.

Fig. 12, a photograph of the oscilloscope display, shows the received pulse train. The display does not

have a flat base line because of the poor low frequency response of the receiving system. The pulse train envelope has in the first portion a more or less irregular structure which eventually transforms into an exponentially decaying curve. The exponential portion indicates that the reiterative wave beam has been established. In Fig. 13(a) and 13(b) the exponential portion has been expanded to illustrate the identical shape of all successive pulses. The original transmitted pulse had an odd characteristic shape which is precisely repeated from iteration to iteration, indicating the absence of any delay distortions. For Fig. 13(a) the reflector plates were spaced 30 meters apart so that the distance between recurring pulses is equal to 60 meters. Therefore, in this photograph, the pulse train corresponds to a total path length of 4.5 km. The attenuation of the pulse train indicates a transmission loss of 1.75 db/km. This is equivalent to a loss of 0.052 db per iteration. The theoretical iteration loss, from Fig. 2 for $(k/D)R^2 = 1.6\pi$, is 0.02 db. Including a reflection loss of about 0.004 db caused by the finite conductivity of the reflector plate, (neglecting coupling loss at the coupling holes) the total calculated iteration loss is $\delta$ 0.024 db or about one half the measured value. This discrepancy, however, is not surprising considering the fact that the radii of curvature of the reflectors were not uniform, and that the surface condition because of difficulty in silver plating was poor. Also, the reflector surfaces could not be polished because of poor adherence of the plated surfaces.

A similar experiment was conducted with the reflector spaced 24.5 meters apart, the curvature of the surface wave substantially deformed to achieve an average radius of curvature of 24.5 meters. The pulse train for this measurement is shown in Fig. 13(b). The measured attenuation in this case corresponds to a transmission loss of 1.2 db/km, or an iteration loss of 0.029 db. The $(k/D)R^2$ value for this case is about $2\pi$. No theoretical diffraction loss figures have been as yet calculated for such high values of $\phi_{max}$ which could be used for quantitative comparison. But this experiment does illustrate that the iteration losses can indeed be made very small. It must be said that the alignment of the reflector plates was not very critical and that no provisions were required for fine adjustment.

The build-up portion of the pulse train depends very much on the curvature of the reflectors. According to theoretical expectations the simultaneous excitation of several beam modes should result in a gradual amplitude increase in the pulse envelope to a maximum, then a gradual transition into the purely exponential decay which occurs when the higher modes have been substantially damped out. Fig. 12 is closest to the theoretical expectations. If the curvature is not uniform or not of the correct radius of curvature the "effective" reflecting surfaces for the various modes do not coincide, that is, the effective path lengths of the various modes are slightly different. This in turn results in different mutual phase shifts for each mode at each reflection, and causes the formation of maxima and minima of the pulse envelope as shown in Fig. 14(a). These maxima and minima are particularly pronounced in Fig. 14(b), where the spacing of the reflector plates was reduced to 21 meters and the reflectors mechanically deformed.

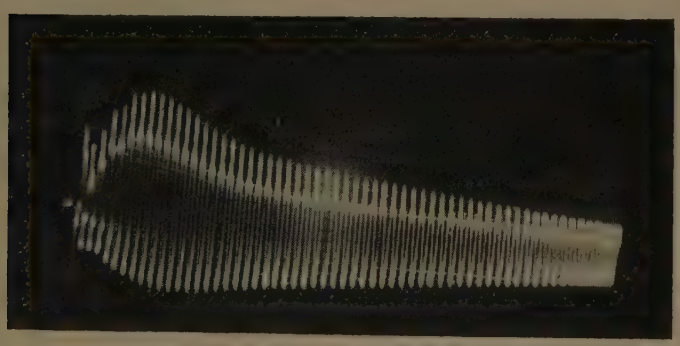
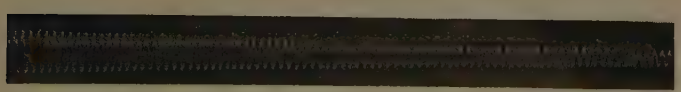


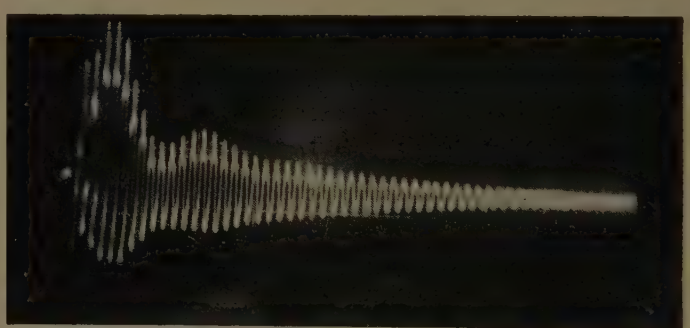
Fig. 12—Envelope of pulses reflected between the reflectors of Fig. 11 for a spacing $D = 30$ meters.

(a)



(b)

Fig. 13—(a) Expanded pulse train of Fig. 12 with $D = 30$ meters; (b) $D = 24.5$ meters.



(a)

(b)

Fig. 14—Build-up of reiterative wave beam from the setup of Fig. 11 for the case of nonuniform reflector curvature.

## V. Conclusion

The experiments support the theory on reiterative wave beams, and demonstrate the feasibility of guiding electromagnetic wave beams with very little loss by resetting the cross-sectional phase distribution at periodic intervals. They also indicate that the diffraction loss can be made small compared to the material losses of the phase transforming devices. The tolerance and alignment of the beam waveguide appear not to be very critical, although more detailed investigations in this respect are warranted.

Since the beams can be deflected by means of reflectors or dielectric wedges, bends in a beam waveguide can readily be performed without introducing any distortions or additional losses. Because of the very rapid decay in the radial direction, the beam is almost completely confined to a cylindrical space of the diameter of the phase-transforming devices and is not affected by obstacles outside this diameter.

## Acknowledgment

The author wishes to thank Dr. F. Schwering for his derivation of the lens loss formulas, Dr. G. Winkler for making the measuring techniques with the Atomicron available, and I. Lerner for his advice on the reflector-plate construction.

# Parasitic Excitation of Circular Antenna Arrays*

T. L. SIMPSON†, member, ire, and J. D. TILLMAN‡, member, ire

*Summary*—**A method of calculating the radiation pattern of a circular antenna array of parallel monopoles or dipoles using symmetrical components is presented. The results obtained when only one element of the array is excited are not particularly good. However, if an additional element is added at the center of the array a fairly good directional pattern results. The sidelobe level is dependent upon the tuning of this central element, and a graphical analysis is given which optimizes this tuning. Calculated and measured patterns are shown and compared. Sidelobe levels of 15 db, and a beamwidth of 78° are typical. The principal advantage of this antenna lies in the fact that the pointing of the beam can be changed by simply changing the element that is excited.**

## Introduction

In recent years, considerable engineering effort has been devoted to the development of antennas capable of radiating directional beams which can be pointed in any desired direction without physical motion in the structure. In such an antenna the beam-steering is accomplished entirely by phase and amplitude control of the excitation currents. Although it is possible to achieve some control over the direction of pointing of a linear array, the angular sector which can be so covered is severely limited. If it is desired to point the beam to any azimuth angle whatever, some configuration involving circular symmetry is indicated.

* Received by the PGAP, January 4, 1960; revised manuscript received, August 4, 1960.

† Developmental Engrg. Corp., Washington, D. C.
‡ University of Tennessee, Knoxville, Tenn.

The circular geometry considered in this regard consists of a number of identical radiators uniformly distributed along the circumference of a circle. If a satisfactory pattern can be obtained at all, then steering the beam could simply be a matter of rotating the excitation accordingly. The array characteristic would be independent of the direction of pointing. With an array having a finite number of elements, an angular rotation of the beam would necessarily be accomplished in discrete steps depending on the number of elements in use.

This paper contains an analysis of circular antenna arrays in which only one element of the array is driven. The configuration which is considered consists of $m$ identical parallel dipoles or monopoles which are spaced uniformly around the circumference of a circle. An additional element, not necessarily having the same half length, is located at the center. When an element in the ring is driven, a pattern with directivity in the plane of the ring results. Considerable control of the sidelobe level can be obtained by tuning the center element. A half-power beamwidth of 80° and a maximum sidelobe level of 14 db results from a typical design.

## Determination of Currents and Patterns

The configuration to be analyzed is shown in Fig. 1. Consider first the situation where all of the elements in the ring are driven by voltages which are symmetrical about the diameter through $\phi = 0$. It can be shown that

such a symmetrical set of $m$ excitation voltages can be resolved into $(m+2)/2$ or $(m+1)/2$ (depending on whether $m$ is even or odd) sequence sets. For simplicity, $m$ will be taken as an even number. The sequence voltages are

$$V_i^{(k)} = V^{(k)} \cos \frac{2\pi}{m} ki, \tag{1}$$

where

$k$ = sequence number, $k = 0, 1, 2 \cdots m/2$,
$i$ = element number, $i = 0, 1, 2 \cdots m/2$,
$V^{(k)}$ = $k$th sequence amplitude.

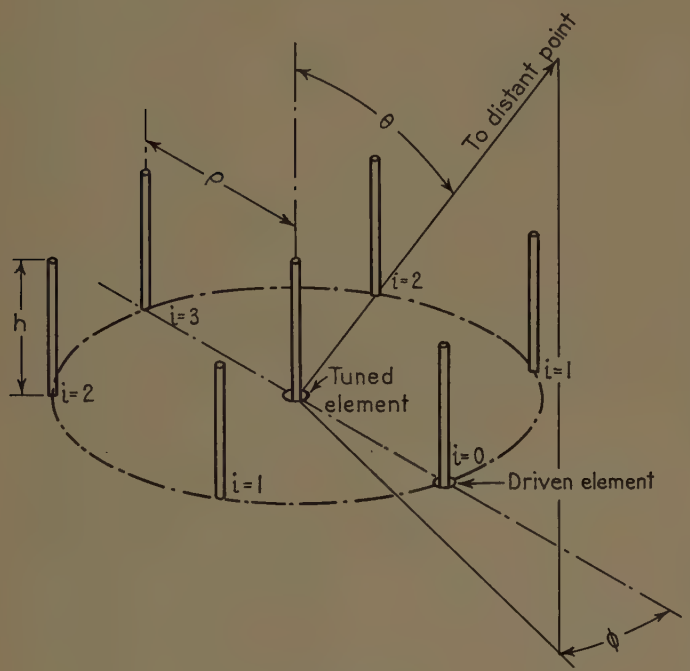



Fig. 1.

We wish to determine $V^{(k)}$ so that

$$V_i = \sum_{k=0}^{m/2} V_i^{(k)} = \sum_{k=0}^{m/2} V^{(k)} \cos \frac{2\pi}{m} ki. \tag{2}$$

Since there are $(m/2)+1$ values of $i$, (2) represents $(m/2)+1$ simultaneous equations which may be solved for the $(m/2)+1$ values of $V^{(k)}$. The result is

$$V^{(k)} = \frac{1}{m} \sum_{i=0}^{m-1} V_i e^{j(2\pi/m)ki}. \tag{3}$$

A similar resolution into sequence components can be carried out for the currents.

The linear equations describing the circuit properties of the array are written in terms of $Z_{ij}$, the mutual impedance between elements $i$ and $j$, and the self impedance. They are

$$V_i = \sum_{j=0}^{m} I_j Z_{ij}. \tag{4}$$

Now if (4) is substituted into (3), for a particular value of $k$, simple algebraic manipulation leads to

$$V^{(k)} = I^{(k)} Z^{(k)}, \tag{5}$$

where

$$Z^{(k)} = \sum_{i=1}^{m} Z_{oi} \cos \frac{2\pi}{m} ki. \tag{6}$$

$Z^{(k)}$ is called the sequence impedance.

It has been shown[1] that the far-field radiation pattern radiated by the $k$th sequence, with phase referred to the array center, is given by

$$F^{(k)} = mI^{(k)} \Big\{ j^k J_k(z) \cos k\phi + \sum_{p=1}^{\infty} [j^{(mp-k)} J_{mp-k}(z) \cos (mp - k)\phi + j^{mp+k} J_{mp+k}(z) \cos (mp + k)\phi] \Big\}, \tag{7}$$

where $z = (2\pi\rho/\lambda) \sin \theta$. The leading term has the form $\cos k\phi$ and is proportional to $J_k(z)$, the $k$th order Bessel function of the first kind. This term is called the principal term, and the remaining terms are called residuals. The largest residual is proportional to $J_{m-k}(z)$, and since, if the order is greater than the argument, the Bessel function is small, the residuals may often be neglected.

Neglecting residuals, the pattern is then

$$F = \sum_{k=0}^{m/2} \frac{V^{(k)}}{Z^{(k)}} j^k J_k(z) \cos k\phi, \tag{8}$$

which is valid for any excitation. If only the element $i = 0$ is driven with voltage $V_o$, the sequence voltages are, from (3),

$$V^{(k)} = \frac{1}{m} V_o \qquad i = 0, m/2$$
$$= \frac{2}{m} V_o \qquad \text{all other } i. \tag{9}$$

Thus, driving one element only excites all the possible sequences. Substituting this into (8) gives

$$F = V_o \frac{J_o(z)}{Z^{(o)}(z)} + 2 \sum_{k=1}^{m/2} j^k \frac{J_k(z)}{Z^{(k)}(z)} \cos k\phi. \tag{10}$$

This includes the first residual for $k = m/2$ which is equal to the principal term. Eq. (10) is general for any number of elements and any radius.

On the basis of (10), extensive calculations were made for arrays of six elements with $2\pi\rho/\lambda$ ranging from 1.5 to 3.0, and with quarter-wave monopoles as elements. Results of spot calculations with a different number of

[1] J. D. Tillman, W. T. Patton, C. E. Blakely, and F. V. Shultz, "The use of a ring array as a skip range antenna," PROC. IRE, vol. 43, pp. 1655–1660; November, 1955.

elements showed no significant differences. Throughout this range, the parasitic excitation resulted in a moderately directional pattern (half-power beamwidths averaged about 80° in azimuth), but was found to produce objectionably high sidelobes, particularly to the rear. Since no way was found to reduce the sidelobe level by varying the available parameters, a modification was sought that would achieve this result without sacrificing the circular symmetry.

The patterns can be improved by including a central element which is not driven. Designating the self-impedance of this element by $Z_{cc}$, the mutual impedance to a ring element as $Z_m$, and the center current by $I_c$, the circuit equations for the array are

$$V_i = I_c Z_m + \sum_{j=0}^{m} I_j Z_{ij}, \tag{11}$$

$$0 = Z_m \sum_{j=0}^{m} I_j + I_c Z_{cc}. \tag{12}$$

Now, when (11) is substituted into (3), it is found that the addition of center element affects only the zero sequence. The new zero-sequence impedance is

$$Z^{(o)\prime} = Z^{(o)} - \frac{Z_m^2}{Z_{cc}}, \tag{13}$$

and the other sequence impedances are unchanged.

The modified azimuthal pattern factor will be as given in (10) but with $Z^{(o)}$ replaced by $Z^{(o)\prime}$ and with the contribution from the center element. The entire pattern is

$$F = V_0 F_e \left\{ \frac{1}{Z^{(o)\prime}} \left[ J_o(z) - \frac{Z_m}{Z_{cc}} \right] + 2 \sum_{k=1}^{m/2} \frac{j^k J_k(z)}{Z^{(k)}} \cos k\phi \right\}. \tag{14}$$

In (14), $F_e$ is the pattern factor of an isolated element of the array.

The self impedance $Z_{cc}$ of the center element may be regarded as the sum of the actual self impedance and an arbitrary tuning reactance, $X_L$. If the center element is identical to the ring elements, then

$$Z_{cc} = Z_{00} + jX_L. \tag{15}$$

A method is now needed to choose $X_L$ to minimize sidelobe level. This will be illustrated for the case of $m=6$. For this case, $Z_m = Z_{01}$. Defining the total omnidirectional part of (25) as $F^{(o)}$, the azimuthal pattern may be written

$$F = F^{(o)} + \sum_{k=1}^{3} F^{(k)}, \tag{16}$$

in which, making use of (14) and (16),

$$F^{(o)} = V_0 F_e \left[ \frac{1}{Z^{(o)}} \left( J_0(z) - \frac{Z_{01}}{Z_{00} + jX_L} \right) \right], \tag{17}$$

and

$$\sum_{k=1}^{3} F^{(k)} = V_0 F_e \left[ 2 \frac{j}{Z^{(1)}} J_1(z) \cos\phi + 2 \frac{j^2}{Z^{(2)}} J_2(z) \cos 2\phi + 2 \frac{j^3}{Z^{(3)}} J_3(z) \cos 3\phi \right]. \tag{18}$$

For convenience in plotting, the expressions of (17) and (18) are multiplied by $10^3$ and normalized to $V_0$.

For a particular choice of circumference the omnidirectional term is a function only of $X_L$ and the directional term is a function only of $\phi$. Symbolically, (17) and (18) become

$$\frac{10^3}{V_0} F^{(o)}(X_L) \text{ and } \frac{10^3}{V_0} \sum_{k=1}^{3} F^{(k)}(\phi),$$

respectively. Thus the omnidirectional component can be controlled independently from the directional part by tuning $X_L$.

To illustrate the optimization technique, consider a six-element ring of vertical $\lambda/4$ monopoles $2.40\lambda$ in circumference. Assume a length-to-radius ratio $h/a = 75$. Fig. 2 presents a complex plane plot of (17) and (18) for these parameters. The omnidirectional term is plotted as the locus of values obtained for $-\infty \leq X_L \leq \infty$ and the directional sum the locus obtained for $0° \leq \phi \leq 180°$; since the excitation is symmetrical the pattern will repeat for $180° \leq \phi \leq 360°$. As the omnidirectional term is independent of $\phi$, it adds directly to the directional term at each angle $\phi$. Thus, in Fig. 2 a particular choice of $X_L$ (and, hence, of $F^{(o)}$) will result in a rigid translation of the directional locus. Construction lines are drawn in the plot to indicate the direction of translation which would maximize the on-axis intensity and to define, also, that translation which would minimize the secondary radiation. The center of the "secondary radiation rectangle," $c$, would be translated to the origin by the addition of its negative, $c'$. Evidently this would minimize the secondary radiation level. In this fortunate case, $c$ is located very nearly opposite the origin from the $\phi = 0°$ point so that translating the directional locus to minimize the secondary level also nearly maximizes the on-axis level. By another fortunate circumstance the $F^{(o)}$ locus passes near the indicated optimal value for $X_L$ in the neighborhood of $-75$ ohms.

In practice, the optimal value for $X_L$ is determined by testing near-optimal values for secondary radiation levels. Fig. 3 shows the azimuthal patterns resulting from three near-optimal values of $X_L$ 0, $-75$, and $-115$ ohms. Evidently $-75$ ohms produces the minimum sidelobe level.

In order to determine pattern variation over a range of frequencies, complex-plane plots of pattern data similar to Fig. 2 were prepared for the arrays whose parameters are tabulated in Table I. A frequency of 300 Mc was chosen for the case shown in Fig. 2. Patterns at 330 Mc and 360 Mc were calculated using the best

value of $X_L$, and beamwidths and sidelobe levels are included.

The usable frequency band extends from slightly below to considerably above 300 Mc. Very poor patterns result at 240 and 270 Mc, with little or no improvement possible by tuning. Above 300 Mc, the pattern deteriorates much less rapidly.

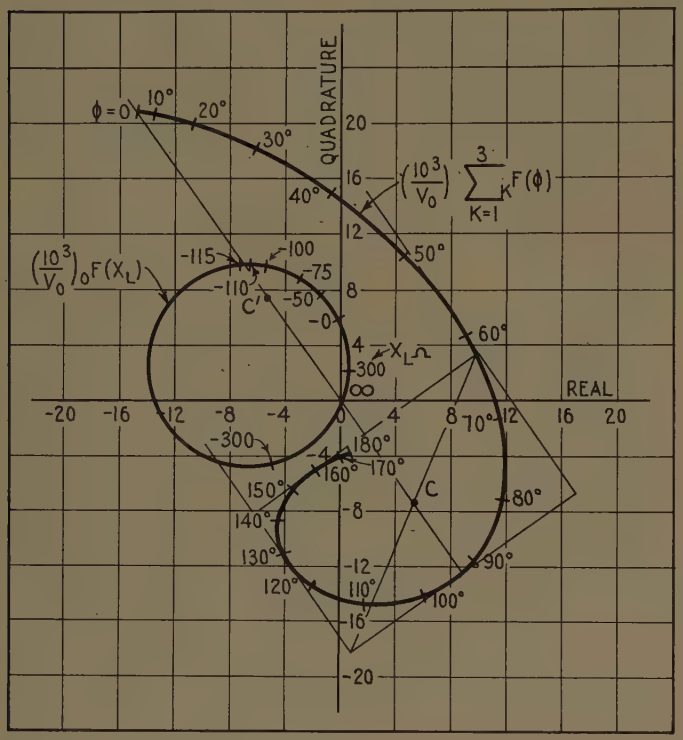



Fig. 2.

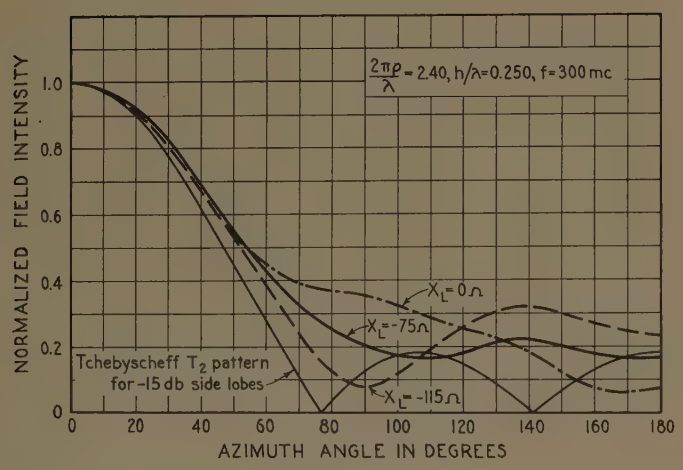



Fig. 3.

TABLE I

| Frequency | Circumference | $X_L$ | Half-Power Beamwidth | Side-Lobe Level |
|---|---|---|---|---|
| 300 Mc | 2.4 λ | −62 | 78° | −13.6 db |
| 330 Mc | 2.64 λ | −163 | 84° | − 9.2 db |
| 360 Mc | 2.88 λ | −180 | 96° | − 8.3 db |

## Model Study

Since the results obtained in the preceding section for the 2.40 λ ring with center tuning were considered promising, this design was tested at a frequency of 300 Mc. The elements were made of quarter-inch bronze rod which resulted in an $h/a$ ratio near 75, the value used in calculations. Both the driven and center elements were constructed by extending the center conductors of 50-ohm coaxial line upward from below the ground plane. Center element tuning was accomplished using a calibrated coaxial tuning stub.

Fig. 4 compares measured with calculated azimuthal patterns of the 2.40 λ ring. Excellent agreement exists in the main beam, although some minor discrepancies appear at the lower levels—a calculated sidelobe of −13.6 db at $\phi = 130°$ was measured as −12.9 db. Fig. 5 compares vertical patterns, measured and calculated, for the array with good agreement.

Experimental investigation at the frequencies of Table I verified the indications of calculated performance. Below about 290 Mc, the pattern deteriorated rapidly regardless of center element tuning. Above the design

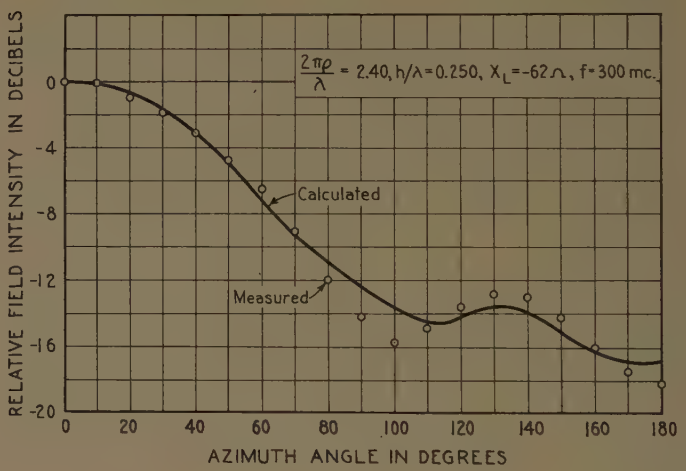



Fig. 4.

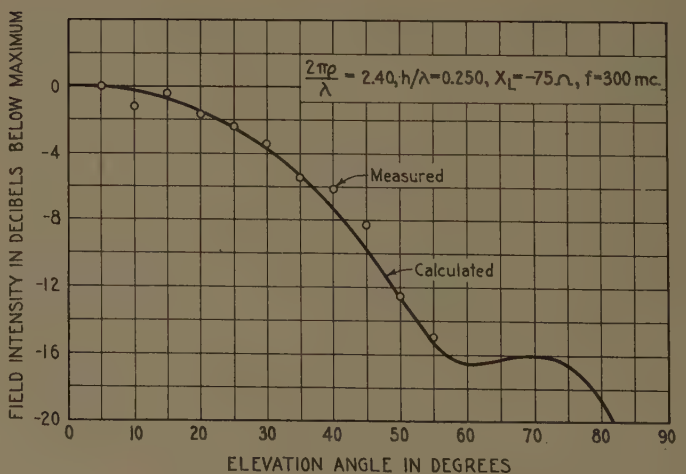



Fig. 5.

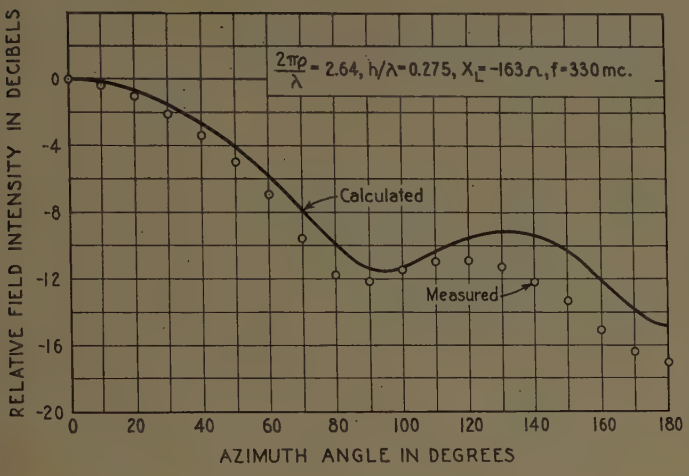



Fig. 6.

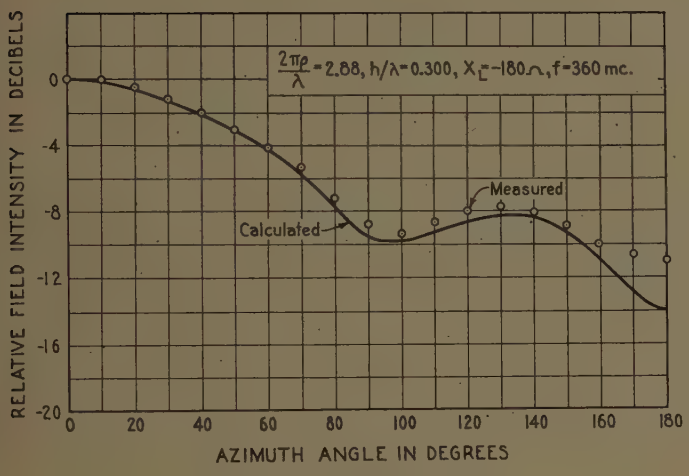



Fig. 7.

frequency of 300 Mc, however, the pattern was much less sensitive to frequency. Figs. 6 and 7 show azimuthal patterns at 300 Mc and 360 Mc, respectively. With increasing frequency the beamwidth becomes wider and sidelobe levels increase.

Impedance measurements were made at the base of the driven ring element. Fig. 8 compares the measured input impedance to the calculated value. While the pointwise correlation is only fair, the order of magnitude and the variation with tuning are in good agreement. Evidently it would be possible to tune out the reactance component of the input impedance by tuning the center element. However, this would disturb the pattern since, at this frequency, a negative tuning reactance is required for optimal sidelobe reduction. Therefore, the input impedance must be tuned separately at the input.

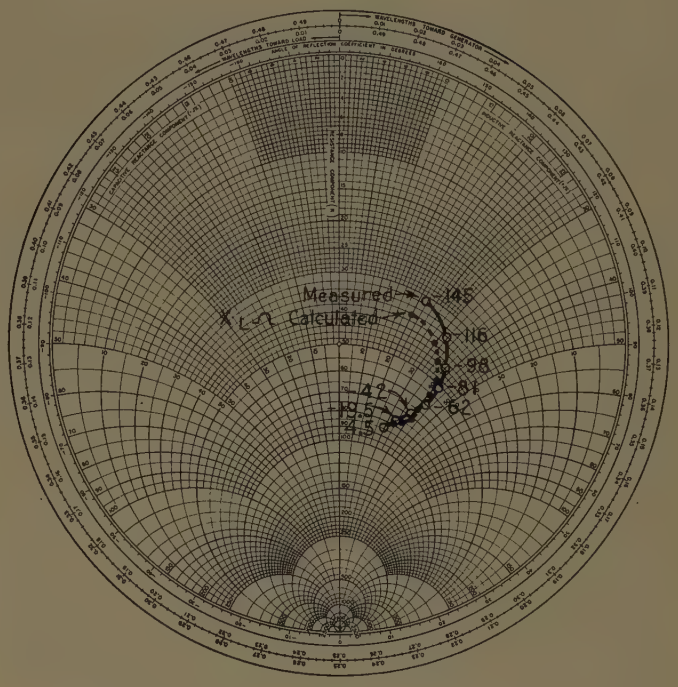



Fig. 8.

## Conclusions

It was found possible to achieve a satisfactory azimuthal pattern with a circular array having six ring elements if a tuned central element were introduced. A graphical analysis was developed whereby an optimal value of tuning reactance could be predicted.

Parasitic excitation imposes definite limitations on the types of patterns obtainable. After the driven element is excited, the currents in the remaining elements are determined by the inherent properties of the array (except for the tuning control on the center element). Such a design is much less powerful, theoretically, than an approach involving a synthesis technique for specifying each elemental current to produce an arbitrary pattern. However, the simplicity of operation of the parasitic array is so attractive that in many cases it is envisioned that pattern requirements might well be relaxed slightly to allow equal competition between driven and parasitic circular arrays.

If single-frequency operation is desired, a very simple switching system is possible to steer the beam. A six-position coaxial switch of the type which short circuits the undesired terminals can be connected to the base of the six elements in the ring through half-wavelength lines. The driven element is then selected by the position of the switch, and the remaining elements are effectively shorted at their base.

# Angular Accuracy of a Phased Array Radar*

L. E. BRENNAN†, MEMBER, IRE

*Summary*—One type of phased-array radar of current interest employs an array of separate receiving elements, each followed by an individual amplifier. These individual signals are combined coherently to form one or more receiving beams for searching, tracking, or performing both functions simultaneously. This paper presents an approach to the theory of angle measurement with a phased array of this type.

In the one-dimensional problem considered here, the receiving antenna consists of a linear array of individual antenna-amplifier elements. The receiver-noise-limited case is considered, in which accuracy is limited by the additive normally distributed noise present in each channel. An expression is derived for the limiting accuracy of angular measurement when a single set of samples is available. This set of samples is obtained simultaneously, one sample from each channel. Next, two methods of implementing the angular measurements are discussed. These are amplitude comparison monopulse and a coherent or phase comparison technique. For large signal-to-noise ratios and for either a square law or a linear envelope detector, the accuracy of amplitude comparison monopulse approaches the theoretical limit. The same accuracy can be achieved with the coherent technique by proper weighting of the individual signals.

## Symbols

$A_1$, $B_1$ = amplitudes of the two quadrature components of $E_1$.

$A_2$, $B_2$ = amplitudes of the two quadrature components of $E_2$.

$$a_1 = \sum_{k=1}^{N} \cos(\phi + k\delta + k\xi).$$

$$a_2 = \sum_{k=1}^{N} \cos(\phi + k\delta - k\xi).$$

$$b_1 = \sum_{k=1}^{N} \sin(\phi + k\delta + k\xi).$$

$$b_2 = \sum_{k=1}^{N} \sin(\phi + k\delta - k\xi).$$

$c$ = an arbitrary constant.

$D_3$ = difference signal; phase comparison, equal weights case.

$D_4$ = difference signal; phase comparison, optimum weights case.

$d$ = spacing between elements of the array.

$E_1$, $E_2$ = signals in the two squinted beams formed for amplitude comparison monopulse.

$e_k$ = signal (voltage or current) in the $k$th channel, includes noise component.

$F(\psi)$ = a function defined in (37).

$k$ = number of the channel (an index).

$L(x_1, y_1, \cdots, x_n, y_n | \delta, \phi)$ = likelihood function for the sample of $x_k$, $y_k$, given $\delta$ and $\phi$.

$N$ = number of channels.

$n_k$ = noise component in $k$th channel.

$S$ = the sum signal (47), used to determine sense of the error signal in phase comparison monopulse.

$\frac{S}{N}$ = signal-to-noise power ratio in one channel.

$t$ = time.

$$U_1 = \sum_{k=1}^{N} u_k \cos k\xi.$$

$$U_2 = \sum_{k=1}^{N} u_k \sin k\xi.$$

$u_k$, $v_k$ = amplitudes of the two quadrature components of $n_k$.

$$V_1 = \sum_{k=1}^{N} v_k \cos k\xi.$$

$$V_2 = \sum_{k=1}^{N} v_k \sin k\xi.$$

$w_k$ = set of weights used in forming difference signal for phase comparison monopulse.

$x_k$, $y_k$ = amplitudes of the two quadrature components of $e_k$.

$\gamma$ = reference phase at center of array.

$\Delta_1$ = difference signal; amplitude comparison monopulse, square law detector.

$\Delta_2$ = difference signal; amplitude comparison monopulse, linear detector.

$\Delta_3$ = difference signal; coherent signal processing, constant weights.

$\Delta_4$ = difference signal; coherent signal processing, optimum weights.

$\delta$ = incremental phase shift between adjacent elements of array due to target displacement from crossover axis.

$\delta^*$ = the estimated value of $\delta$.

$\theta$ = angle between axis of array and line of sight to target.

$\theta'$ = angle between crossover axis and line of sight to target.

$\lambda$ = wavelength.

$\xi$ = incremental phase shift between adjacent elements of array to provide monopulse squint angles.

* Received by the PGAP, August 19, 1960.
† The RAND Corp., Santa Monica, Calif.

$\rho_1$, $\rho_2$ = envelopes of signals in the two channels corresponding to squinted beams for amplitude comparison monopulse.
$\sigma$ = rms noise (voltage or current) in the individual channels.
$\sigma_{\Delta 1}$ = rms error in $\Delta_1$ due to noise.
$\sigma_{\Delta 2}$ = rms error in $\Delta_2$ due to noise.
$\sigma_{\Delta 3}$ = rms error in $\Delta_3$ due to noise.
$\sigma_{\Delta 4}$ = rms error in $\Delta_4$ due to noise.
$\sigma_{\delta 1}$ = rms error in $\delta$; amplitude comparison monopulse, square law detector.
$\sigma_{\delta 2}$ = rms error in $\delta$; amplitude comparison monopulse, linear detector.
$\sigma_{\delta 3}$ = rms error in $\delta$; coherent signal processing, constant weights.
$\sigma_{\delta 4}$ = rms error in $\delta$; coherent signal processing, optimum weights.
$\sigma_{\delta}^*$ = lower bound for the rms error in $\delta$.
$\phi$ = reference phase of signal.
$\psi = N\xi$.
$\omega = 2\pi \cdot$ signal frequency.

## I. Introduction

THERE is increasing interest today in phased array radars which use a set of separate antenna elements, each followed by an individual amplifier, in place of a more conventional receiving antenna. These individual signals can be combined coherently to form several receiving beams simultaneously. The theory of angle measurements with this type of array is discussed here. For simplicity, only the one-dimensional problem is considered, and the antenna elements are assumed to be equally spaced in a linear array.

In practice, the angular accuracy of this type of radar can be limited by any of several effects, including receiver noise, external noise, component phase or amplitude errors, and atmospheric refraction. The following analysis considers the errors due to receiver noise, *i.e.*, noise originating in the individual antenna-amplifier channels.

## II. Statement of the Problem

A linear array of equally spaced antenna elements is assumed, each followed by an individual amplifier, as illustrated in Fig. 1. When a plane wave is incident at an angle $\theta$ from the axis of the linear array, the individual amplifier outputs (voltage or current) can be expressed in the form

$$e_k = \cos(\omega t + \phi + k\delta) + n_k, \tag{1}$$

where

$e_k$ = signal in $k$th channel,
$\omega = 2\pi \cdot$ signal frequency,
$\phi$ = reference phase of signal,
$\delta$ = incremental phase shift between channels,
$n_k$ = noise in $k$th channel.

The incremental phase shift, $\delta$, is related to the angle of arrival by

$$\delta = \frac{2\pi d}{\lambda} \cos\theta, \tag{2}$$

where

$\lambda$ = wave length,
$d$ = spacing between elements,
$\theta$ = angle of arrival measured from axis of array.

For convenience, the following analysis considers errors in the incremental phase shift $\delta$. These can be converted readily to errors in angle of arrival $\theta$, using (2).

It is assumed that the reference phase of the signal $\phi$ is unknown, as is generally the case in radar systems. The effects of equipment errors, either random or systematic, are not considered here, nor are errors in normalizing the signals in the individual channels.

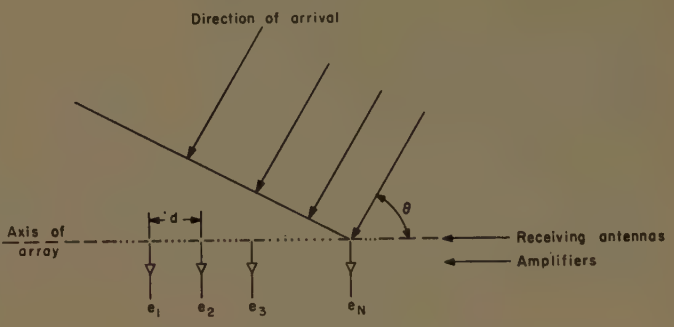



Fig. 1—Linear phased array geometry.

The noise components in the individual channels are assumed to be independent and normally distributed. It will be convenient to express them in the form

$$n_k = u_k \cos\omega t + v_k \sin\omega t, \tag{3}$$

where

$u_k$, $v_k$ = quadrature components of noise in the $k$th channel,

and

$$\overline{u_k} = \overline{v_k} = 0$$

$$\overline{u_k^2} = \overline{v_k^2} = \overline{n_k^2} = \sigma^2. \tag{4}$$

The bar above a quantity is used to denote average value. The signal-to-noise power ratio in each individual channel is then

$$\frac{S}{N} = \frac{1}{2\sigma^2}. \tag{5}$$

In terms of the above definitions, the specific problem considered here is the following: Given a set of signals $e_k$, obtained by sampling the outputs of the $N$ channels of a linear array simultaneously, how accurately can the incremental phase shift $\delta$ be estimated? Since a single set of $e_k$ samples is assumed, this corresponds to the single-pulse accuracy of a pulsed radar.

In Section III, a lower bound is obtained for the rms error in estimating $\delta$. Sections IV and V consider two specific methods of estimating $\delta$. These are amplitude comparison monopulse and a coherent phase comparison technique.

## III. Lower Bound on Error in $\delta$

The accuracy with which the incremental phase shift $\delta$ can be estimated from one sample of the set of signals $e_k$ is limited by noise. In this section, a theorem of statistics is used to obtain a lower bound on the rms error in the estimate of $\delta$. This theorem applies specifically to all regular unbiased estimates. Denoting the estimated value of $\delta$ by $\delta^*$, an unbiased estimate is one for which the expectation of $\delta^*$ is equal to $\delta$, *i.e.*,

$$E(\delta^*) = \delta. \tag{6}$$

The conditions for regularity are more subtle and are discussed by Cramér.[1]

It is convenient to express the signals, $e_k$, in the form

$$e_k = x_k \cos \omega t - y_k \sin \omega t \tag{7}$$

where $x_k$ and $y_k$ are the two quadrature components of one sample from the $k$th channel. The likelihood function for the sample $\{x_k, y_k\}$ is then

$$L(x_1, y_1, \cdots, x_N, y_N \mid \delta, \phi) = \frac{1}{(2\pi\sigma^2)^N} \prod_1^N \cdot \exp\left\{ - \frac{[x_k - \cos(k\delta + \phi)]^2 + [y_k - \sin(k\delta + \phi)]^2}{2\sigma^2} \right\}. \tag{8}$$

This expression follows directly from the assumption that the noise components in the individual channels are independent and normally distributed.

Although we are concerned only with estimating $\delta$, there are two unknown parameters in this case, $\delta$ and $\phi$, so the expression for a joint estimate must be used. From Cramér,[1] the mean square error for all regular unbiased estimates of $\delta$ has the lower bound

$$\sigma_{\delta^*}^2 \geq \frac{E\left\{\left(\frac{\partial \log L}{\partial \phi}\right)^2\right\}}{E\left\{\left(\frac{\partial \log L}{\partial \phi}\right)^2\right\} \cdot E\left\{\left(\frac{\partial \log L}{\partial \delta}\right)^2\right\} - \left[E\left\{\frac{\partial \log L}{\partial \phi} \cdot \frac{\partial \log L}{\partial \delta}\right\}\right]^2}. \tag{9}$$

From (8)

$$\frac{\partial(\log L)}{\partial \phi} = -\frac{1}{\sigma^2} \sum_1^N \{[x_k - \cos(k\delta + \phi)] \sin(k\delta + \phi) - [y_k - \sin(k\delta + \phi)] \cos(k\delta + \phi)\} \tag{10}$$

and

$$\frac{\partial(\log L)}{\partial \delta} = -\frac{1}{\sigma^2} \sum_1^N k\{[x_k - \cos(k\delta + \phi)] \sin(k\delta + \phi) - [y_k - \sin(k\delta + \phi)] \cos(k\delta + \phi)\}. \tag{11}$$

Since the noise components in the individual channels are independent of each other, as are the two quadrature components in each channel, we have

$$E\{[x_k - \cos(k\delta + \phi)][y_l - \sin(l\delta + \phi)]\} = 0 \text{ all } k, l \tag{12}$$

and

$$E\{[x_k - \cos(k\delta + \phi)][x_l - \cos(l\delta + \phi)]\} = E\{[y_k - \sin(k\delta + \phi)][y_l - \sin(l\delta + \phi)]\} = 0 \quad k \neq l = \sigma^2 \quad k = l. \tag{13}$$

Combining (9)–(13) gives the following expression for the lower bound on mean square error in $\delta$:

$$\sigma_{\delta^*}^2 \geq \frac{12\sigma^2}{N^3 - N}. \tag{14}$$

## IV. Amplitude Comparison Monopulse

In amplitude comparison monopulse, two beams are formed, as illustrated in Fig. 2. Each beam is formed by summing all $N$ outputs with the appropriate relative phases. The two resulting signals are envelope-detected and the difference of the two envelopes is calibrated in terms of angle from cross-over $\theta'$, as shown in Fig. 2. The approximate incremental phase shift (or angle of incidence) must be known *a priori* to permit formation of two beams which are pointed roughly in the direction of incidence. A square law envelope detector is assumed in Sections IV-A through IV-C, and the linear detector case is discussed in Section IV-D. In both cases a large signal-to-noise ratio is assumed so that noise cross-modulation terms can be neglected.

[1] Harold Cramér, "Mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., ch. 32; 1946.

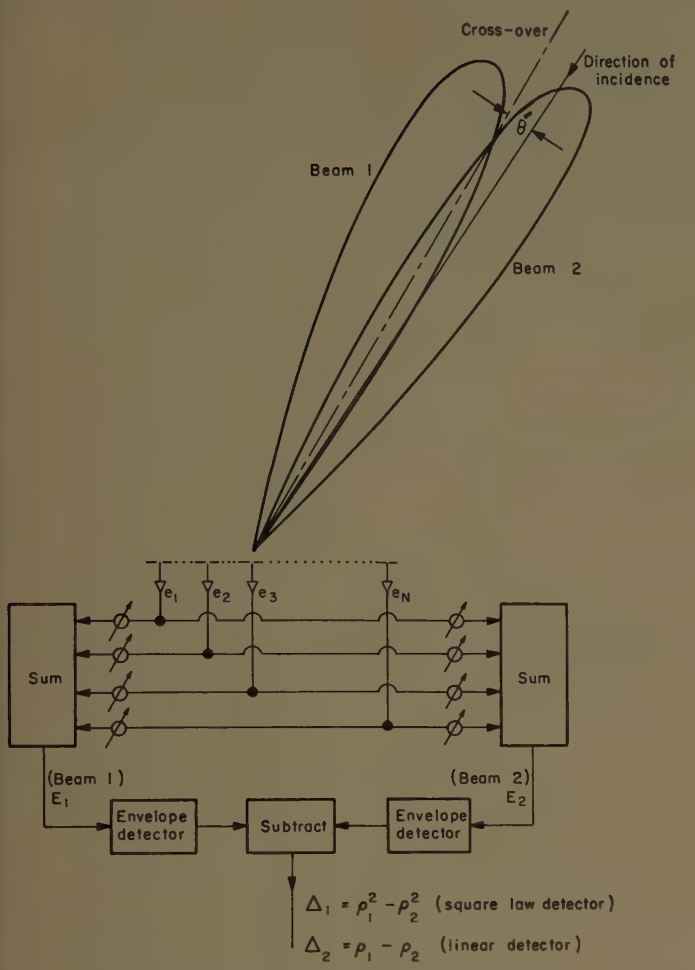



Fig. 2—Amplitude comparison monopulse.

In terms of the individual channel voltages, $e_k$, the sum voltages for the two beams are

$$E_1 = \sum_1^N e_k/\underline{k\xi}$$

$$= \sum_1^N [\cos(\omega t + \phi + k\delta + k\xi) + u_k \cos(\omega t + k\xi) + v_k \sin(\omega t + k\xi)],$$

$$E_2 = \sum_1^N e_k/\underline{-k\xi}$$

$$= \sum_1^N [\cos(\omega t + \phi + k\delta - k\xi) + u_k \cos(\omega t - k\xi) + v_k \sin(\omega t - k\xi)], \quad (15)$$

where

$\delta$ = incremental phase shift due to displacement of cross-over from incidence angle ($\theta'$ in Fig. 2)

$\xi$ = incremental phase shift for beam squinting.

Since the noise voltages in the two channels are correlated and will partially cancel in the difference signal $\Delta$, it is important to retain information on the relative noise phases.

The sum signals can also be expressed in the following form:

$$E_1 = A_1 \cos \omega t + B_1 \sin \omega t$$

$$E_2 = A_2 \cos \omega t + B_2 \sin \omega t \quad (16)$$

where

$$A_1 = a_1 + U_1 + V_2$$

$$B_1 = -b_1 - U_2 + V_1$$

$$A_2 = a_2 + U_1 - V_2$$

$$B_2 = -b_2 + U_2 + V_1 \quad (17)$$

and

$$a_1 = \sum_1^N \cos(\phi + k\delta + k\xi)$$

$$a_2 = \sum_1^N \cos(\phi + k\delta - k\xi)$$

$$b_1 = \sum_1^N \sin(\phi + k\delta + k\xi)$$

$$b_2 = \sum_1^N \sin(\phi + k\delta - k\xi)$$

$$U_1 = \sum_1^N u_k \cos k\xi$$

$$U_2 = \sum_1^N u_k \sin k\xi$$

$$V_1 = \sum_1^N v_k \cos k\xi \quad (18)$$

$$V_2 = \sum_1^N v_k \sin k\xi.$$

In terms of these quantities, the outputs of the two square-law envelope detectors are

$$\rho_1^2 = A_1^2 + B_1^2$$

$$\rho_2^2 = A_2^2 + B_2^2. \quad (19)$$

Finally, the error signal, which is denoted by $\Delta_1$ in this case, is

$$\Delta_1 = \rho_1^2 - \rho_2^2. \quad (20)$$

To obtain the error in the measurement of incremental phase shift $\delta$, two quantities must be calculated: the rms value of $\Delta_1$ due to noise $\sigma_{\Delta_1}$, and the error slope or derivative of $\Delta_1$ with respect to $\delta$. The rms error in estimating $\delta$ is then

$$\sigma_{\delta 1} = \frac{\sigma_{\Delta_1}}{\left(\frac{\partial_{\Delta_1}}{\partial \delta}\right)}. \quad (21)$$

*A. The Error Slope—Square Law Detector*

First, we will obtain an expression for the error slope in the absence of noise. In this case, the difference signal is

$$\Delta_1 = \rho_1^2 - \rho_2^2 = (a_1^2 + b_1^2) - (a_2^2 + b_2^2). \quad (22)$$

Substituting the sums of (18) for $a_1$, $a_2$, $b_1$, and $b_2$ into this expression gives

$$\Delta_1 = \frac{\sin^2 \frac{N}{2}(\delta + \xi)}{\sin^2\left(\frac{\delta + \xi}{2}\right)} - \frac{\sin^2 \frac{N}{2}(\delta - \xi)}{\sin^2\left(\frac{\delta - \xi}{2}\right)}. \quad (23)$$

The following identities were used in summing the series of (18):

$$\sum_1^N \sin k\alpha = \frac{\cos\frac{\alpha}{2} - \cos\left(N + \frac{1}{2}\right)\alpha}{2 \sin \frac{\alpha}{2}},$$

$$\sum_1^N \cos k\alpha = \frac{\sin\left(N + \frac{1}{2}\right)\alpha - \sin\frac{\alpha}{2}}{2 \sin \frac{\alpha}{2}}. \quad (24)$$

Differentiating the expression for error signal $\Delta_1$ of (23) gives the following value for error slope at crossover:

$$\left.\frac{\partial \Delta_1}{\partial \delta}\right|_{\delta=0} = \frac{N \sin (N\xi) \sin^2\left(\frac{\xi}{2}\right) - \sin(\xi) \sin^2\left(\frac{N\xi}{2}\right)}{\sin^4\left(\frac{\xi}{2}\right)}. \quad (25)$$

*B. Derivation of $\sigma_{\Delta_1}$—Square Law Detector*

Next we obtain an equation for the mean square error in $\Delta_1$ due to noise, *i.e.*,

$$\sigma_{\Delta_1}^2 = \overline{(\Delta_1 - \bar{\Delta}_1)^2}. \quad (26)$$

From (17), (19), and (20)

$$\Delta_1 - \bar{\Delta}_1 = 2[U_1(a_1 - a_2) + V_2(a_1 + a_2) + U_2(b_1 + b_2) + V_1(b_2 - b_1)] + 4(U_1V_2 - U_2V_1). \quad (27)$$

For the large signal-to-noise ratio case considered here, the second term of (27), which consists of noise cross-modulation products, can be neglected. The equation for $\sigma_{\Delta_1}^2$ is then

$$\sigma_{\Delta_1}^2 = 4[(a_1 - a_2)^2\overline{U_1^2} + (a_1 + a_2)^2\overline{V_2^2} + (b_1 + b_2)^2\overline{U_2^2} + (b_2 - b_1)^2\overline{V_1^2} + 2(a_1 - a_2)(b_1 + b_2)\overline{U_1U_2} + 2(a_1 + a_2)(b_2 - b_1)\overline{V_1V_2}]. \quad (28)$$

Terms of the form $\overline{U_iV_j}$ are zero since the two quadrature components of noise in the individual channels are independent.

Next, consider the individual quantities occurring in (28). The quantity $\overline{U_1^2}$ can be expressed as follows:

$$\overline{U_1^2} = \overline{\left(\sum_1^N u_k \cos k\xi\right)\left(\sum_1^N u_l \cos l\xi\right)} \quad (28)$$

$$= \overline{\sum_1^N u_k^2 \cos^2 k\xi}$$

$$= \sigma^2 \sum_1^N \cos^2 k\xi$$

$$= \frac{\sigma^2}{4}\left[(2N + 1) + \frac{\sin(2N + 1)\xi}{\sin \xi}\right]. \quad (29)$$

Since the noise components in the different channels are independent, terms of the form $u_ku_l$ in (29) were replaced with

$$\overline{u_ku_l} = 0 \qquad k \neq l$$

$$\overline{u_k^2} = \sigma^2 \qquad \text{all } k. \quad (30)$$

The corresponding expressions for the other terms in (28) involving $U_i$ or $V_i$ are

$$\overline{U_2^2} = \overline{V_2^2} = \frac{\sigma^2}{4}\left[(2N + 1) - \frac{\sin(2N + 1)\xi}{\sin \xi}\right] \quad (31)$$

$$\overline{V_1^2} = \overline{U_1^2} \quad (32)$$

and

$$\overline{U_1U_2} = \overline{V_1V_2} = \frac{\sigma^2}{4}\left[\frac{\cos \xi - \cos(2N + 1)\xi}{\sin \xi}\right]. \quad (33)$$

Terms of the form $(a_1 \pm a_2)$ or $(b_1 \pm b_2)$ in (28) are functions of $\delta$. For simplicity, only the $\delta = 0$ case was considered here, which would be closely approximated in many cases with closed-loop tracking systems. Substituting the expressions of (18) for $a_1$, $b_1$, $a_2$, and $b_2$ into (28) gives

$$\sigma_{\Delta_1}^2 = 16\left[\overline{U_1^2}\left(\sum_1^N \sin k\xi\right)^2 + \overline{U_2^2}\left(\sum_1^N \cos k\xi\right)^2 - 2\overline{U_1U_2}\left(\sum_1^N \sin k\xi\right)\left(\sum_1^N \cos k\xi\right)\right] \quad (34)$$

which after substitution of (29), (31), (32), (33) and further simplification becomes

$$\sigma_{\Delta_1}^2 = 2\sigma^2\left[\frac{4N \sin^2 \frac{N\xi}{2} \sin \xi + \sin 2N\xi - 2 \sin N\xi}{\sin \xi \sin^2 \frac{\xi}{2}}\right]. \quad (35)$$

### C. Error in Estimating δ—Square Law Detector

Expressions have now been obtained for the error slope [see (25)] and the mean square error in $\Delta_1$ due to noise [see (35)]. These can be combined using (21) to obtain the rms error in measuring incremental phase shift $\delta$. The incremental phase shift $\xi$, which determines the monopulse squint angle has been retained as a parameter, and it is interesting to see how the error in $\delta$ varies with this quantity. Combining (21), (25), and (35), the mean square error in $\delta$ can be expressed as follows when $N$ is large:

$$\sigma_{\delta 1}^2 = \frac{\sigma_{\Delta_1}^2}{\left(\frac{\partial \Delta_1}{\partial \delta}\right)_{\delta=0}^2} \doteq \frac{\sigma^2}{2N^3} F(\psi) \tag{36}$$

where

$$\psi = N\xi$$

and

$$F(\psi) = \psi^3 \frac{\left(4\psi \sin^2 \frac{\psi}{2} + \sin 2\psi - 2 \sin \psi\right)}{\left(\psi \sin \psi - 4 \sin^2 \frac{\psi}{2}\right)^2}. \tag{37}$$

The function $F(\psi)$ is plotted in Fig. 3. For $\psi = \pi$, the squint angle is 1/2 beamwidth, *i.e.*, the centers of the two monopulse beams are separated by one beamwidth.

Note that the function $F(\psi)$ is constant for small $\psi$; *i.e.*, the accuracy of angular measurement is independent of squint angle in this region. This is explained by noting that as the squint angle decreases, the error slope decreases, while the correlation between the noise components in the two beams increases. The noise therefore cancels more completely in the difference signal. The error slope and noise in the difference signal $\Delta$ decrease proportionately so that the resulting error in $\delta$ remains constant.

For squint angles less than 1/2 beamwidth (*i.e.*, $\psi = \pi$), the value of $F(\psi)$ is 24, and the corresponding rms error in measurement of $\delta$ is

$$\sigma_{\delta 1} \doteq \frac{\sqrt{12}\sigma}{N^{3/2}}. \tag{38}$$

This equation is valid for large $N$, and in this case the rms error in $\delta$ approaches the theoretical limit discussed in Section III. When $N$ is small, the mean square error in $\delta$ approaches the limit given in (14) for small squint angles (*i.e.*, a term of the form $N^3 - N$ occurs in the denominator).

### D. Linear Envelope Detector

The corresponding analysis for the case of a linear envelope detector will be outlined here. Again a large signal-to-noise ratio is assumed and the rms error in the difference signal is computed only for the case of $\delta = 0$.

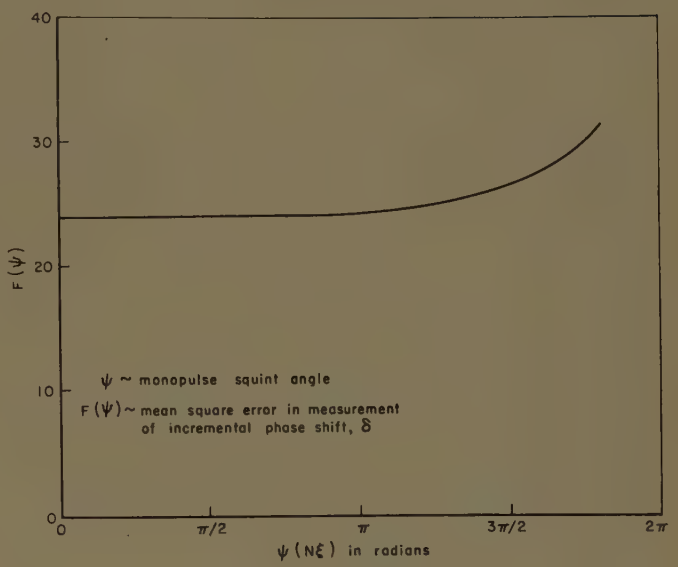



Fig. 3—The function $F(\psi)$.

With a linear envelope detector, the difference signal, which is denoted by $\Delta_2$ in this case, is

$$\Delta_2 = \rho_1 - \rho_2 \tag{39}$$

where $\rho_1$ and $\rho_2$ are defined as before [(15)–(19)].

In the absence of noise, the difference signal is

$$\Delta_2 = \frac{\sin \frac{N(\delta+\xi)}{2}}{\sin \frac{\delta+\xi}{2}} - \frac{\sin \frac{N(\delta-\xi)}{2}}{\sin \frac{\delta-\xi}{2}}. \tag{40}$$

Differentiating this expression to obtain the error slope gives

$$\left.\frac{\partial \Delta_2}{\partial \delta}\right|_{\delta=0} = \frac{N \sin \frac{\xi}{2} \cos \frac{N\xi}{2} - \sin \frac{N\xi}{2} \cos \frac{\xi}{2}}{\sin^2 \frac{\xi}{2}}. \tag{41}$$

In addition to the error slope, one must also evaluate the rms error in $\Delta_2$ due to noise. Using (39), (15), (19), and noting that for the large signal-to-noise ratio case cross-modulation products of noise can be neglected, gives

$$\Delta_2 = \sqrt{a_1^2 + b_1^2 + 2a_1(U_1 + V_2) + 2b_1(U_2 - V_1)} - \sqrt{a_2^2 + b_2^2 + 2a_2(U_1 - V_2) - 2b_2(U_2 + V_1)}. \tag{42}$$

For the case of $\delta = 0$ and large signal-to-noise ratio this equation reduces to

$$\Delta_2 - \overline{\Delta_2} = \frac{U_1(a_1 - a_2) + V_1(b_2 - b_1) + U_2(b_1 + b_2) + V_2(a_1 + a_2)}{\sqrt{a_1^2 + b_1^2}}. \tag{43}$$

As before, this expression is square and averaged to obtain $\sigma_{\Delta_2}{}^2$.

Again the rms error in $\delta$, denoted by $\sigma_{\delta 2}$ in this case, is the ratio of $\sigma_{\Delta_2}$ and the error slope. The same end result for rms error in $\delta$ is obtained in this case, namely, for the case of large signal-to-noise ratio and a linear detector

$$\sigma_{\delta 2}{}^2 = \frac{\sigma^2}{2N^3} F(\psi) \tag{44}$$

where $F(\psi)$ is defined in (37).

## V. Coherent Methods of Angle Measurement

Two coherent or linear signal-processing techniques are discussed in this section. In both cases, a difference signal proportional to the incremental phase shift $\delta$, is formed by combining the individual signals $e_k$, coherently (*i.e.*, at RF or IF where phase is preserved). In the first case, the signals from each half of the array are summed individually with equal weights. The difference between these two partial sums is proportional to the incremental phase shift $\delta$. This is essentially a phase comparison monopulse technique, the difference signal being proportional to the difference in phase of the two partial sums. The second technique, discussed in Section V-B, is similar, but weighs the different signals $e_k$, unequally.

### A. Constant Weights

First consider the case where the signals are weighted equally and the difference signal has the form

$$D_3 = \sum_{N/2+1}^{N} e_k - \sum_{1}^{N/2} e_k. \tag{45}$$

For convenience, we assume the array contains an even number of elements, $N$. Again, the approximate angle of incidence must be known to steer the beams formed by the two halves of the array. From (1) and (45), the signal component of the difference signal is given by

$$D_3 = \left\{ -2 \frac{\sin^2 \dfrac{N\delta}{4}}{\sin \delta/2} \right\} \sin (\omega t + \gamma) = \Delta_3 \sin (\omega t + \gamma), \tag{46}$$

where

$$\gamma = \phi + \frac{N+1}{2}\delta = \text{reference phase at center of array.}$$

The amplitude of the difference signal $\Delta_3$ is a function of the magnitude of the incremental phase shift $\delta$, while the sense of the error signal is determined by the RF phase of $D_3$. In this case, a sum signal must also be formed to determine the sense of the error signal

$$S = \sum_{1}^{N} e_k = K \cos (\omega t + \gamma). \tag{47}$$

For small pointing errors, *i.e.*, small values of $\delta$, the quantity $K$ is approximately equal to the number of channels, $N$. Of course, the sum signal $S$ also contains a noise term, so that at very low signal-to-noise ratios, errors in sign of the error signal may occur frequently. Again we will consider only the large signal-to-noise ratio case, and assume that the amplitude of the sin $(\omega t+\gamma)$ component of $D_3$ is obtained coherently.

In this case, the rms noise in the difference signal $\Delta_3$ is

$$\sigma_{\Delta 3} = \sqrt{N}\,\sigma. \tag{48}$$

The error slope in this case is

$$\left(\frac{\partial \Delta_3}{\partial \delta}\right) = \frac{\sin^2 \dfrac{N\delta}{4} \cos \delta/2 - \dfrac{N}{2} \sin \dfrac{\delta}{2} \sin \dfrac{N\delta}{2}}{\sin^2 \delta/2}. \tag{49}$$

For small values of $\delta$, this reduces to

$$\left(\frac{\partial \Delta_3}{\partial \delta}\right)_{\delta=0} = -\frac{N^2}{4}. \tag{50}$$

The rms error in $\delta$ is again obtained by dividing the rms error in the difference signal by the error slope, *i.e.*,

$$\sigma_{\delta 3} = \frac{\sigma_{\Delta_3}}{\left(\dfrac{\partial \Delta_3}{\partial \delta}\right)_{\delta=0}} = \frac{4\sigma}{N^{3/2}}. \tag{51}$$

The limiting accuracy of (14) is not achieved in this case when $N$ is greater than 2. For large $N$, the rms error obtained with the coherent technique described here is 13 per cent greater than the error for amplitude comparison monopulse.

### B. Variable Weights[2]

In the preceding subsection, it was shown that when the signals are all weighted equally, the limiting accuracy (see [14]) is not achieved. From the following heuristic argument, one would expect an improvement in accuracy when a set of unequal weights is used. The error signal obtained by subtracting the two end signals from a multi-element array $(e_N - e_1)$ is greater for a given $\delta$ than the error signal from the two middle elements of the array. However, the rms noise in these two difference signals is the same. Hence, a set of variable weights which favor the end elements of an array should improve the ratio of error-signal to noise. In the following paragraphs it will be shown that this is indeed the case, and that with an optimum set of weights the limiting accuracy is achieved.

[2] The possibility of improving accuracy with a set of variable weights was suggested by D. L. Margerum, of Systems Labs. Corp.

In the general case of variable weights, the difference signal $D_4$ is

$$D_4 = \sum_1^{N/2} w_k[e_{N-k+1} - e_k]$$

$$= \left(\sum_1^{N/2} 2w_k \sin\left(k - \frac{N+1}{2}\right)\delta\right) \sin(\omega t + \gamma) \qquad (52)$$

$$= \Delta_4 \sin(\omega t + \gamma).$$

Again, $N$ is assumed to be an even number, and the large signal-to-noise ratio case is considered. As before, the sum signal is used to determine the sense of the error signal and to extract coherently only one quadrature component of signal plus noise.

The mean square noise in the error signal $\Delta_4$ is then

$$\sigma_{\Delta_4}^2 = 2\sigma^2 \sum_1^{N/2} w_k^2. \qquad (53)$$

Again considering the case of small $\delta$, the error slope obtained by differentiating $\Delta_4$ of (52) is given by

$$\frac{\partial \Delta_4}{\partial \delta} = 2 \sum_1^{N/2} w_k \left(k - \frac{N+1}{2}\right). \qquad (54)$$

Next, one would like to find the set of weights $w_k$ which minimizes the ratio of $\sigma_{\Delta 4}$ to the error slope, *i.e.*, the set for which

$$\frac{\partial}{\partial w_k}\left[\frac{\partial \Delta_4}{\left(\frac{\partial \Delta_4}{\partial \delta}\right)}\right] = 0 \qquad k = 1, 2, \cdots, N/2. \qquad (55)$$

This optimum set of weights is proportional to distance from the center of the array, *i.e.*,

$$w_k = c\left(k - \frac{N+1}{2}\right) \qquad (56)$$

where $c$ is a constant.

From (53), (54), and (56), the corresponding mean square error in $\delta$ is

$$(\sigma_{\delta 4})^2 = \frac{12\sigma^2}{N^3 - N}. \qquad (57)$$

This is the same mean square error as was obtained in Section III for the limiting accuracy and again in Section IV for amplitude comparison monopulse.

## VI. Conclusions

A lower bound was obtained for the rms error in angle measurement with a linear phased array for the single sample, receiver-noise-limited case. It was shown that this limiting accuracy is achieved with amplitude comparison monopulse for the more tractable case of large signal-to-noise ratios. It was also shown that with amplitude comparison monopulse, angular accuracy is independent of squint angle over a wide range of squint angles—from zero to approximately 1/2 beamwidth.

When the individual signals are processed coherently, as discussed in Section V, and weighted equally, the angular error exceeds the theoretical limit by 13 per cent. However, when an optimum set of weights is used, the limiting accuracy is again achieved for large signal-to-noise ratios. It was shown that optimum weighting in this case implies weights proportional to distance from the center of the array.

Some of the results obtained here are also valid at low signal-to-noise ratios. The expression for limiting accuracy derived in Section III is valid at all signal-to-noise ratios. However, the results for amplitude comparison monopulse are invalid at low signal-to-noise ratio, since the noise cross-modulation terms in the envelope detector output were neglected. Noise cross-modulation would predominate at low signal-to-noise ratios, resulting in rapid degradation of angular accuracy. The analysis of Section V for the coherent signal processing case is valid at low signal-to-noise ratios, provided an accurate reference or sum signal is available. However, at low signal-to-noise ratios the reference signal would also contain a large noise component, resulting in frequent errors in the sign of the error signal.

When closed-loop angle-tracking is used, the accuracy would exceed that estimated here for the single sample case. In effect, with closed-loop tracking many samples are integrated to obtain a current estimate of angle. Although the closed-loop case was not considered explicitly, the results obtained here provide an essential input for the analysis of linear arrays employing this technique.

# A Spiral-Doublet Scanning Array*

JOHN R. DONNELLAN†

*Summary*—**An array of eight spiral doublets, each doublet consisting of a pair of equally-excited two-wire spirals in the same plane and wound in opposite senses, has been used to scan a beam of horizontal polarization over a ±40° range simply by rotations of the spiral doublets. The combination of a microstrip ring network and twin lead provided a balanced feed for each spiral in the array. Sidelobe levels were those predicted (25 db) for the broadside pattern and—at the test frequency of 1430 Mc—were, in general, below 19 db over the 80° scan. A novel microstrip feed harness provided isolation between the doublets and served to absorb most circulating currents. This printed circuit feedboard consisted of judiciously-grouped microstrip ring networks and provided the desired feeding coefficients to within ±0.1 db. An application of the array for tracking purposes is also described.**

## Introduction

IT has been shown[1,2] that two equally excited spirals of opposite sense, placed side by side in the same plane—a spiral doublet—radiate a combined field which is everywhere linearly polarized. The direction of polarization is controlled by rotating one spiral of the doublet relative to the other. The phase of the radiated field is changed, while holding direction of polarization constant, by rotating both elements of the doublet equal amounts but in opposite directions.

A beam scan of linear polarization from an eight doublet array of spirals has previously been obtained.[2,3] The direction and width of each main beam was in very close agreement with theoretical predictions, but sidelobes were excessive. Moreover, the beam scanned successfully 25° more to one side than to the other. Each spiral in this array had been excited from a coaxial line, *i.e.*, an unbalanced feed. The spiral on the other hand is a symmetrical device which indicates the need for a balanced input. Further reduction of sidelobe levels and over-all improvement of performance of the array appeared to depend to a large extent upon improvement of individual spiral performance. Also, an improved method for obtaining correct phasing between spiral doublets was deemed necessary. With these needs in mind, it was decided that a new and improved array would be constructed.

* Received by the PGAP, July 1, 1960; revised manuscript received, October 3, 1960.

† Microwave Antennas and Components Branch, Electronics Div., U. S. Naval Res. Lab., Washington, D. C.

[1] J. A. Kaiser, "Scanning Arrays Using the Flat Spiral Antenna," Naval Res. Lab., Washington, D. C., Rept. No. 5103; March, 1958.

[2] J. A. Kaiser, "The Archimedean two-wire spiral antenna," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-8, pp. 312–323; May, 1960.

[3] J. A. Kaiser, "Spiral Antennas Applied to Scanning Arrays," presented at Electronic Scanning Symp., AF Cambridge Res. Ctr., Cambridge, Mass.; April, 1958.

## Spirals and Associated Components

Fig. 1 depicts the geometry of the Archimedean spiral used in the experimental work described in this report. A spiral doublet, therefore, consists of two spirals in the same plane, one wound in the sense shown in Fig. 1, the other wound in the opposite sense. An array of such doublets leads to an antenna of an arbitrary linear polarization which can be made to scan by changing the relative phase between the spiral doublets.

A front view of the eight-spiral doublet array used in the experimental work is shown in Fig. 2. The spacing of the spirals above the ground plane is 0.25 λ and the spacing between doublets (and also between two spirals

Fig. 1—Archimedean spiral antenna.

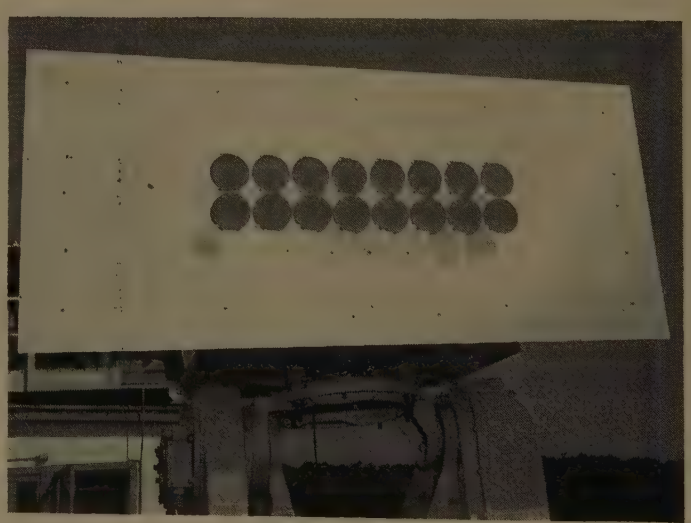


Fig. 2—Eight-spiral doublet array.

of each doublet) is 0.5 λ (10.5 cm) at 1430 Mc, which was the test frequency. All spirals along the top row are of one sense while those along the bottom row are of the opposite sense. This array was intended to scan only in the horizontal plane with an arbitrary linear polarization; horizontal polarization was chosen for the initial experimental work.

Each spiral was fed with 150-ohm twin lead, one end of the twin lead being connected to the input terminals at the center of the spiral while the other end was connected to a microstrip ring network (Fig. 3). Attaching a two-wire line at points B and C assures that the input terminals of the spiral are fed 180° out of phase, a condition necessary for optimum spiral performance. Thus the rat-race provided a transition (balun) from an unbalanced line to a balanced line, and the combination of ring network and twin lead provided a balanced feed for the spiral antennas. This balanced feed eliminated the "squint" which had been associated with the coaxially fed spirals.

Aquadag was applied to the areas between the filaments of the last quarter-turns of each spiral in order to absorb any currents which reached the outside edges of the spirals. Since the Aquadag was located beyond the area from which radiation occurred, there was no effect on the currents reaching the radiating region from the input terminals. The Aquadag served to attenuate currents which, if allowed to be reflected from the ends of the spiral, would radiate in a sense opposite to that of the currents traveling from the input terminals. These two opposite-sense currents would result in an elliptically polarized radiated field from the spiral, the degree of ellipticity being a function of the relative magnitudes of the two currents. Thus Aquadag was employed to improve the axial ratio of the field radiated from each spiral. An axial ratio of 1.5 db with no Aquadag painted on the spirals was reduced to 0.5 db when Aquadag was applied.

## Feed Structure

The novel feed harness used to excite the array is diagrammed in Fig. 4. It was composed of judiciously grouped microstrip rat-races of the same general construction as the balun of Fig. 3. The board was carefully designed to give the proper Tchebycheff illumination—into matched loads—for 25-db sidelobe levels. An analysis of design considerations follows.

Power is fed into the center rat-race at point 1 (Fig. 4) and divides equally at points 2 and 4, the currents at these points being 180° out of phase. The initial hybrid network is thus used to divide the input power into two equal amounts, one half going to one side of the feedboard, the other half going to the opposite side (a mirror image). If there are any reflected currents at points 5 and 6, these currents will arrive back at points 2 and 4 in phase and will be absorbed by the matched-load terminating arm 3. For convenience of description, the board has been divided into eight sections, I to VIII. Consider section IV. Power entering this section at point 6 divides equally at points 7 and 8, the currents at these points being 180° out of phase; these currents then continue on to points 10 and 11. If line length $l'$ is equal to line length $l''$, no power will enter arm 12, but all will arrive at point 13. If $l' = l'' \pm 0.5\,\lambda$, all the power will enter arm 12, none arriving at point 13. If $l' = l'' \pm 0.25\,\lambda$, the power divides equally between 12 and 13. Therefore, any desired power distribution between points 12 and 13 can be obtained by proper choice of $|l' - l''|$. The power distribution for values of $|l' - l''|$ ranging from 0 to 0.5 λ in steps of 0.05 λ was measured at 1430 Mc:

$$|l' - l''| = \left(\frac{n}{10}\right)\frac{\lambda}{2}, \qquad n = 0, 1, 2, \cdots, 10.$$

The normalized curve of the experimental results is shown in Fig. 5 along with the theoretical curve. The maximum error is 5 per cent. The experimental values were used in the design of the feedboard. With this information, any desired fraction of the power could be tapped off at point 12 (Fig. 4), the remaining power continuing on from point 13 into section III. Here, again by proper choice of line lengths, the desired frac-

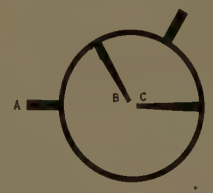



Fig. 3—Microstrip ring network balun.

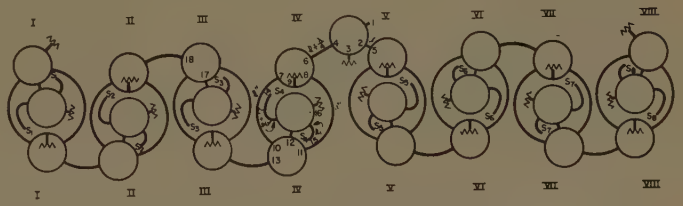



Fig. 4—Microstrip feedboard design.

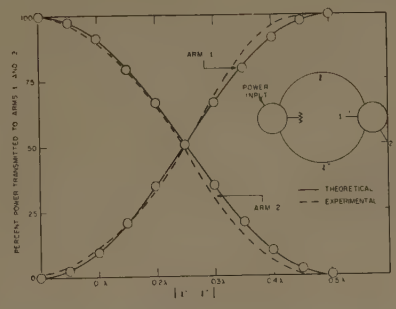



Fig. 5—Relative power transmission (theoretical and experimental) vs path difference $|l' - l''|$ for scheme shown.

tion of the power could be tapped off at point 17, the remainder continuing on from point 18 into section II, etc. In this manner, the proper feeding coefficients for the array could be obtained.

Again, consider section IV of the feedboard. Power entering the center rat-race at point 12 divides equally between arms 14 and 15. Coaxial cables were attached at these points in order to feed the spiral doublet. Any reflections at points 14 and 15 (or beyond these points, down the transmission line to the spiral antennas themselves) will be absorbed by the matched-load terminating arm 16. A similar argument holds for the other seven doublets. The output arms of the feedboard used to feed doublets 1 to 8 are indicated in Fig. 4 by $S_1$ to $S_8$ respectively. A similar extended analysis will show that this feedboard provides isolation between the doublets and serves to absorb most of the circulating currents by series of dumpings into matched loads.

## Experimental Results

The spiral array was phased by using it as the receiver while a pyramidal horn oriented for horizontal polarization was used as the transmitter. The transmitter and receiver were separated by a distance of 72 feet (approximately $3\times 2d^2/\lambda$). A bolometer detector was connected to the output arm of the feedboard (arm 1 in Fig. 4) and then by coaxial cable to an amplifier. With the receiver directly facing the transmitter, all doublets in the array were first set for horizontal polarization. The next step was to rotate all doublets by the proper amount to obtain an on-axis in-phase condition. When the array had been phased, the pattern in the horizontal plane was recorded (Fig. 6). The pattern covers the range from $+108°$ to $-108°$. Sidelobe levels obtained (over a $\pm 90°$ range) are those predicted. The positions of the lobes are also in close agreement with the predicted positions, while the 3-db main beamwidth (14°) is two degrees narrower than the predicted width (16°).

It had previously been determined that in order to direct the beam to some angle $\theta$, the phase difference between adjacent doublets had to be equal to approximately $3\theta$ (more accurately, 180 sin $\theta$); this $3\theta$ relationship holds for scan angles up to approximately 45°. Thus Fig. 7, in addition to the broadside pattern, shows the $\pm 10°$ beam displacements obtained by introducing a phase difference of $\pm 30°$ between adjacent doublets. In one case, the upper spirals of doublets 5–8 were rotated clockwise 15°, 45°, 75°, and 105°, respectively, while the upper spirals of doublets 4, 3, 2, and 1 were rotated counter-clockwise 15°, 45°, 75°, and 105°, respectively; in the other case, the rotation directions were reversed. Rotation of the upper spiral of a doublet in one direction through an angle $\alpha$ causes a corresponding rotation of the lower spiral in the opposite direction through an equal angle $\alpha$. Thus constant polarization was maintained while only the phase between doublets was changed. In this manner, the phase center was balanced

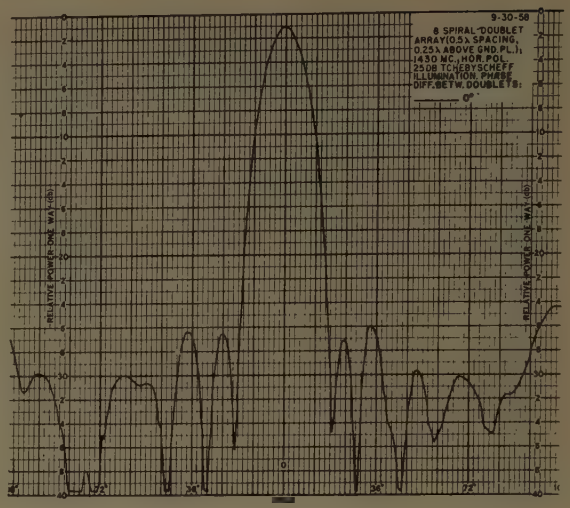



Fig. 6—Experimental broadside pattern

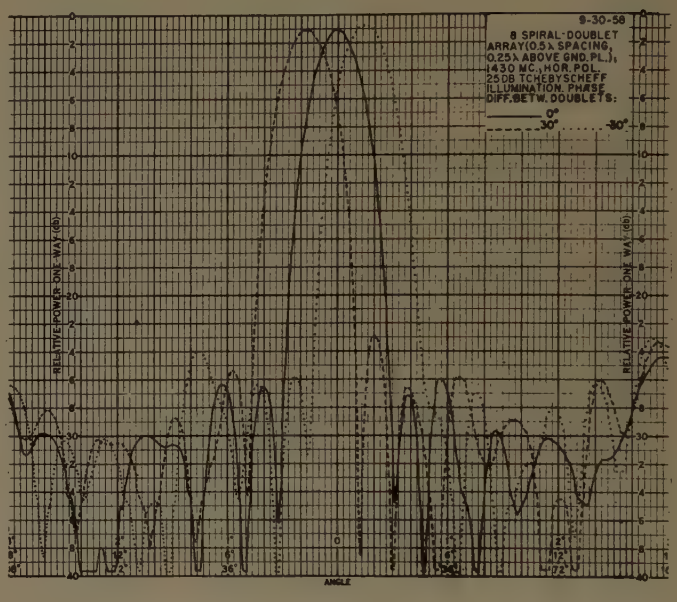



Fig. 7—Patterns produced by phase differences between doublets of 0° and $\pm 30°$ (main beam directed to 0° and $\pm 10°$, respectively).

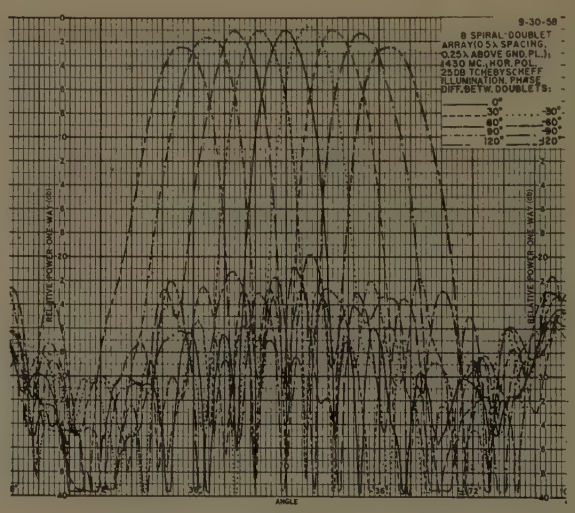



Fig. 8—Patterns produced by phase differences between doublets of 0°, $\pm 30°$, $\pm 60°$, $\pm 90°$, and $\pm 120°$ (main beam directed to 0°, $\pm 10°$, $\pm 20°$, $\pm 30°$, and $\pm 40°$, respectively).

about the center of the array, and scanning was achieved by simple proper rotations of the spiral doublets. In a similar manner, beam displacements of ±20°, ±30°, and ±40° were obtained as indicated in Fig. 8.

Table I shows the variation of maximum sidelobe level (over a ±90° range relative to the broadside position) and 3-db main beamwidth with angle of scan (0° to ±40°). In general, the sidelobes are increasing with increasing scan; the average sidelobe levels, however, are considerably below the figures given in the table (*e.g.*, all sidelobes other than the maximum at scan angles ±40° are greater than 19 db down from the main beam). The 3-db main beamwidths are, in general, 2° narrower than the theoretically predicted beamwidths, with both increasing with increasing scan angles.

TABLE I

VARIATION OF MAXIMUM SIDELOBE LEVEL AND 3-DB MAIN BEAMWIDTH WITH SCAN ANGLE

| Scan Angle (Degrees) | Maximum Sidelobe Level (db) | 3-db Main Beamwidth (Degrees) |
|---|---|---|
| 0 | 25.0 | 14.0 |
| ±10 | 22.0 | 14.5 |
| ±20 | 20.0 | 14.5 |
| ±30 | 20.0 | 16.5 |
| ±40 | 17.5 | 18.0 |

At scan angles greater than ±40°, sidelobes were excessive (*e.g.*, the maximum sidelobe level at ±45° was 16 db down from the main beam). The cross-polarized broadside pattern showed lobes 20 db down from the main beam.

An additional feature of this array can be seen in Fig. 9. Shown here again is the broadside pattern as obtained with the detector at arm 1 (see Fig. 4) of the feedboard, as described previously. If the detector and the load-terminating arm 3 are switched, the sharp null shown in Fig. 9 is obtained. This null occurs because a 180° phase shift is delivered to one side of the feedboard when the detector and load are switched, so that the four doublets on one side of the feedboard are 180° out of phase with the four doublets on the other side. The pattern thus obtained would be useful for tracking purposes.

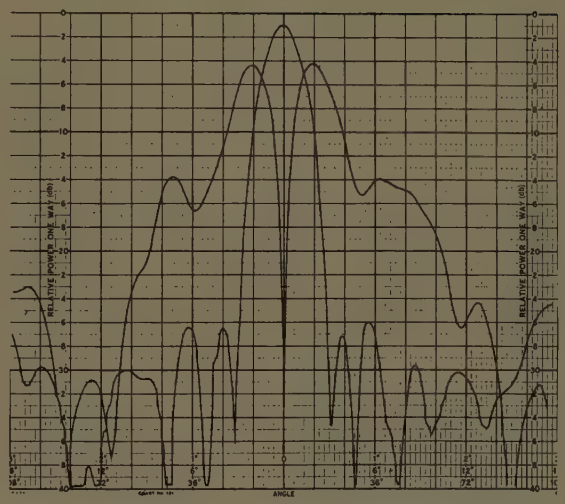


Fig. 9—A broadside pattern and a split-beam pattern resulting from a 180° phase shift delivered to one side of the array.

ACKNOWLEDGMENT

The author wishes to express his appreciation to Dr. A. E. Marston for his original ideas and interest, to J. A. Kaiser for several helpful suggestions, to F. L. Hennessey for the technique of phasing the array, and to R. J. Wiegand for his excellent technical assistance.

# Leaky Wave Antennas II: Circular Waveguides*

L. O. GOLDSTONE†, SENIOR MEMBER, IRE, AND A. A. OLINER†, FELLOW, IRE

*Summary*—**The propagation characteristics of leaky waves on slotted circular cylinders are derived by the methods described in an earlier paper on leaky rectangular waveguides. The cross section of a typical leaky circular waveguide is represented by a radial transmission line network incorporating a lumped terminating admittance which characterizes the slot discontinuity and the external region. The expression for the lumped admittance is obtained by combining an integral equation solution with a variational procedure. The resonances of this transverse network, which yield the complex propagation constants for the leaky waves, are solved by perturbation techniques to produce results in simple and practical form. Solutions are obtained for leaky waves corresponding to two different excitations of a slotted cylinder and these results are shown to compare favorably with measured values.**

## I. Introduction

FLUSH-MOUNTED antennas of the type classed as leaky waveguides radiate all along the length of the waveguide and can be characterized by a complex propagation constant along the waveguide. The characteristics of leaky wave antennas and their analysis by microwave network techniques have been discussed in an earlier paper, where applications were made to leaky waveguides of *rectangular* cross section.[1] It is the purpose of this paper to present the application of such network methods to the analysis of slotted *circular* waveguides of the type shown in Fig. 1.

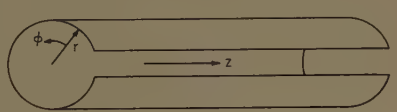



Fig. 1—Slotted circular waveguide.

The representation of a slotted circular waveguide by a transverse equivalent network involves a nonuniform transmission line of the so-called *radial* type. The modes required for such a radial transmission line representation are $E$-type and $H$-type modes which propagate in the radial direction but are distinguished by vanishing longitudinal (axial) components of the magnetic field and electric field, respectively. These modes are similar to the modes employed in connection with rectangular leaky waveguides[1] in that no coupling between the two types can occur in longitudinally uniform structures. A discussion of these modes is presented in Section II.

The parameters of the lumped circuit portion of the transverse equivalent network which characterizes the slot discontinuity and the external region are obtained by a combination of small-aperture results, which are derived in the Appendix, and a variational procedure. The Appendix treats the diffraction by the slot of a radial wave arising from within the slotted cylinder, for two different exciting fields. The diffraction problems are phrased in terms of integral equations for the slot electric field; solutions for the slot fields and for the slot magnetic polarizabilities are obtained in the static small-aperture limit. These polarizabilities are then used in Section III in the evaluation of the slot discontinuity susceptances, while the slot fields are employed in variational expressions for the radiation conductances.

The calculation of the leaky wave propagation constants is accomplished by employing a perturbation technique for the solution of the transverse resonance equation which is similar to that used for rectangular leaky waveguides.[1]

Two types of leaky waves in slotted circular waveguides are considered in this paper. These are the $E$-type and $H$-type leaky waves which can be regarded as perturbations of the $E_{01}$ and $H_{11}$ modes, respectively, in a closed round waveguide. The theoretical results agree very well with measured values for both types of leaky waves when the slot is narrow. The agreement is not as good for very wide slots because a perturbation procedure was used and because the lumped susceptances occurring in the transverse networks are evaluated in the small-aperture limit. The problem of calculating the propagation constants of these two leaky waves has been treated previously by Harrington.[2] Although there exist similarities between his method and the network procedure described here, there are also significant differences between them. The advantages of the network procedure have been discussed in the previous paper.[1] It might, however, be mentioned that the network treatment results in expressions for the propagation constant which are much simpler to compute from than those of Harrington and which, furthermore, are explicit in form.

As mentioned above, the diffraction problems treated in the Appendix also yield the magnetic polarizabilities of an infinite slit in an infinite cylinder for both the longitudinal and transverse components of the magnetic field. It is interesting to compare these results with those for an infinite slit in an infinite conducting plane. The

* Received by the PGAP, July 7, 1960. The research described in this paper was performed at the Microwave Res. Inst. of the Polytechnic Inst. of Brooklyn under Contract No. AF 19(604)-2031 with the AF Cambridge Res. Ctr.

† Microwave Res. Inst., Polytechnic Inst. of Brooklyn, Brooklyn, N. Y.

[1] L. O. Goldstone and A. A. Oliner, "Leaky wave antennas I: rectangular waveguides," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-7, pp. 307–319; October, 1959.

[2] R. F. Harrington, "Propagation along a slotted cylinder," *J. Appl. Phys.*, vol. 24, pp. 1366–1371; November, 1953.

results in the literature for the plane are available only for the case for which there is no field variation along the slit. Under these conditions, the transverse magnetic polarizabilities, which correspond to weak excitation of the slits, are identical when the width of the slit in the plane is replaced by the arclength of the slit in the cylinder. For the longitudinal magnetic polarizability, which corresponds to strong excitation of the slits, the results for the plane and the cylinder are slightly different from each other. Further details are included in Section III.

## II. Radial Transmission Line Modes

Although the cross sections transverse to the radial direction in circular cylindrical coordinates are nonuniform, it is nevertheless possible to describe the fields transverse to the radial direction in terms of mode functions which possess orthogonality properties over the transverse cylindrical surfaces. In terms of such modes, circular cylindrical waveguide structures may be represented by transverse, radial transmission lines.

The field expansion employed here differs from that presented in the "Waveguide Handbook."[3] The mode function concept, as discussed there, requires that the mode functions be independent of the coordinate along the transmission direction, with the result that a vector modal representation is not possible for radial geometry. However, if the mode function concept is modified to permit a variation of the mode functions with $r$, it becomes possible to derive a vector modal representation with orthogonality properties over the cylindrical $\phi z$ cross section. Such a representation is discussed below for the case where there is an arbitrary exponential variation in the axial ($z$) direction.

Using such radial mode functions, a representation of the transverse (to $r$) vector fields may be written as[4]

$$\boldsymbol{E}_t(r,\phi) = \sum_n V_n'(r)\boldsymbol{e}_n'(r,\phi) + V_n''(r)\boldsymbol{e}_n''(r,\phi), \quad (1a)$$

$$\boldsymbol{H}_t(r,\phi) = \sum_n I_n'(r)\boldsymbol{h}_n'(r,\phi) + I_n''(r)\boldsymbol{h}_n''(r,\phi). \quad (1b)$$

The primes in the above representation denote the $E$-type modes for which $H_z=0$, and the double primes denote the $H$-type modes for which $E_z=0$.

A set of mode functions appropriate for radial transmission line representations has been discussed elsewhere by Altschuler and Goldstone.[5] The radial transmission line equations, characteristic impedances and radial propagation wavenumbers given in this reference are the same as those to be employed here. However, it should be noted that the mode functions which are to be used in this paper differ from those given in the reference cited in that the $\phi$ components $e_{n\phi}$ and $h_{n\phi}$ of the mode functions used here contain a factor of $1/r$ which does not appear in the corresponding functions given in the above reference. The mode functions as defined here possess orthogonality properties over the cylindrical surfaces transverse to $r$ and are normalized as follows:

$$\int_0^{2\pi} \boldsymbol{h}_n'^* \times \boldsymbol{r}_0 \cdot \boldsymbol{e}_m' r d\phi = \int_0^{2\pi} \boldsymbol{h}_n''^* \times \boldsymbol{r}_0 \cdot \boldsymbol{e}_m'' r d\phi = \delta_{mn}, \quad (2)$$

$$\int_0^{2\pi} \boldsymbol{h}_n'^* \times \boldsymbol{r}_0 \cdot \boldsymbol{e}_m'' r d\phi = \int_0^{2\pi} \boldsymbol{h}_n''^* \times \boldsymbol{r}_0 \cdot \boldsymbol{e}_m' r d\phi = 0, \quad (3)$$

where $\boldsymbol{r}_0$ is a unit vector in the radial direction and the asterisk denotes the complex conjugate.

Since both the cases which are treated in this paper involve longitudinal field components which are even functions of the angular coordinate $\phi$, the required mode functions become:

$E$-type modes:

$$e_{zn}' = \sqrt{\frac{\epsilon_n}{\pi}} \cos n\phi, \quad (4a)$$

$$h_{zn}' = 0, \quad (4b)$$

$$h_{\phi n}' = -\sqrt{\frac{\epsilon_n}{\pi}} \frac{\cos n\phi}{r}, \quad (4c)$$

$$e_{\phi n}' = \frac{jk_z}{k^2 - k_z^2} \frac{n}{r} \frac{\sin n\phi}{\sqrt{\pi}}, \quad (4d)$$

$H$-type modes:

$$h_{zn}'' = \sqrt{\frac{\epsilon_n}{\pi}} \cos n\phi, \quad (5a)$$

$$e_{zn}'' = 0, \quad (5b)$$

$$e_{\phi n}'' = \sqrt{\frac{\epsilon_n}{\pi}} \frac{\cos n\phi}{r}, \quad (5c)$$

$$h_{\phi n}'' = \frac{jk_z}{k^2 - k_z^2} \frac{n}{r} \frac{\sin n\phi}{\sqrt{\pi}}, \quad (5d)$$

where

$$n = 0, 1, 2, \cdots$$

and

$$\epsilon_n = \begin{cases} 1/2, & n = 0 \\ 1, & n > 0. \end{cases}$$

The orthogonality of the mode functions given in (4) and (5) is easily verified by substitution into (3).

[3] N. Marcuvitz, "Waveguide Handbook," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 10, Section 1.7; 1951.

[4] The exponential $z$ dependence has been suppressed.

[5] H. M. Altschuler and L. O. Goldstone, "On network representations of certain obstacles in waveguide regions," IRE Trans. on Microwave Theory and Techniques, vol. MTT-7, pp. 216–217; April, 1959.

## III. Transverse Equivalent Networks

The two "lowest" leaky waves in slotted round waveguide can be regarded as perturbations of the $H_{11}$ and $E_{01}$ modes in a closed guide. Each of these cases poses a different transverse discontinuity problem for the evaluation of the lumped circuit parameters which characterize the slot and the external region.

The susceptances which characterize the slot are in each case obtained by a "small-aperture" procedure using results which are derived in the Appendix. The conductances are obtained by the application of a variational procedure employing small-aperture quasi-static trial fields.

### A. Perturbed $H_{11}$ Mode

The coordinate system and the electric field distribution for this case are indicated in Fig. 2. The appropriate

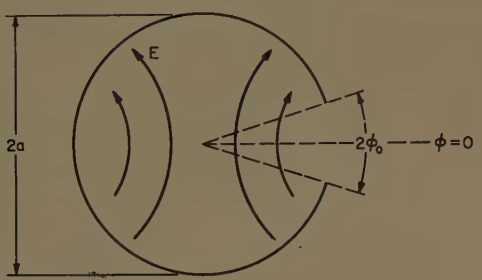



Fig. 2—Slotted circular waveguide carrying the perturbed $H_{11}$ mode.

discontinuity problem is that of the $n=1$ radial $H$-type mode [in (5)] incident on the slot from within the waveguide. From the solution of this discontinuity problem, given in the Appendix, one obtains

$$\int_{-\phi_0}^{\phi_0} \boldsymbol{r}_0 \times \boldsymbol{E}_t a d\phi = \frac{M_z}{j\omega\epsilon} I_1'' \boldsymbol{h}_1''(a, 0), \tag{6}$$

where

$\boldsymbol{E}_t$ is the tangential electric field in the slot,

$I_1''$ is the discontinuity in the modal current of the $n=1$ radial $H$-type mode at the slot,

$\boldsymbol{h}_1''(a, 0)$ is the value of the $n=1$ $H$-type mode function at the center of the slot, and

$M_z = \dfrac{\pi}{2} \dfrac{k^2}{(k^2 - k_z^2)} \dfrac{1}{\ln(2/\phi_0)}$ is the magnetic polarizability of the slot for the axial component of the magnetic field.[6]

[6] The axial magnetic polarizability of an infinite slit in an infinite conducting plane is known for the case $k_z=0$ to be

$$M_z = \frac{\pi}{2} \frac{1}{\ln(8/k\gamma d)}.$$

(See N. Marcuvitz, "Waveguide Circuit Theory: Coupling of Waveguides by Small Apertures," Polytechnic Inst. of Brooklyn, Brooklyn, N. Y., Rept. R-157-47, PIB-106, p. 42, eq. (26c); 1947. It is seen that this polarizability is similar in form to that for the cylinder, but, in contrast to the cylinder result, it is dependent on frequency.

The modal voltage of the $n=1$ mode at the slot is given by

$$V_1 = \int_{-\phi_0}^{\phi_0} \boldsymbol{h}_1''^*(a, \phi) \cdot \boldsymbol{r}_0 \times \boldsymbol{E}_t(\phi) a d\phi$$

$$\approx \boldsymbol{h}_1''^*(a, 0) \cdot \int_{-\phi_0}^{\phi_0} \boldsymbol{r}_0 \times \boldsymbol{E}_t a d\phi. \tag{7}$$

The substitution of (6) into (7) then yields

$$V_1'' \approx \frac{M_z}{j\omega\epsilon} I_1'' \boldsymbol{h}_1''(a, 0) \cdot \boldsymbol{h}_1''^*(a, 0). \tag{8}$$

When the polarizability value and the appropriate mode functions from (5) are substituted into (8), the unnormalized slot susceptance is obtained directly as

$$\frac{I_1''}{V_1''} = jB = j\omega\epsilon \left(\frac{k^2 - k_z^2}{k^2}\right) 2 \ln (2/\phi_0), \tag{9}$$

while the normalized susceptance becomes

$$B' = \frac{B}{Y_0''} = 2\kappa a \ln (2/\phi_0), \tag{10}$$

where

$$Y_0'' = \frac{\kappa}{\omega\mu a}$$

and

$$\kappa = \sqrt{k^2 - k_z^2}.$$

The susceptance is normalized to the characteristic admittance of the angularly symmetric ($n=0$) $H$-type mode rather than the $n=1$ mode. This is a matter of convenience, since any normalization is satisfactory provided all admittances are normalized in the same way.

The conductance portion of the equivalent circuit which terminates the radial transmission line in this case can be calculated from

$$G = \frac{P}{|V_1''|^2}, \tag{11}$$

where $P$ is the radial real power flow per unit length and $V_1''$ is the mode voltage of the $n=1$ radial mode at the slot. The power $P$ is given by

$$P = \text{Re} \int_{-\phi_0}^{\phi_0} \boldsymbol{E} \times \boldsymbol{H}^* \cdot \mathbf{r}_0 a d\phi = \text{Re} \int_{-\phi_0}^{\phi_0} E_\phi H_z^* a d\phi. \tag{12}$$

The magnetic field $H_z$ can be represented by

$$H_z = \sum_{n=0}^{\infty} I_n'' h_{zn}'', \tag{13}$$

with

$$I_n'' = \vec{Y}_n'' V_n'' \tag{14a}$$

and

$$V_n'' = \int_{-\phi_0}^{\phi_0} \boldsymbol{h}_n''^* \cdot \boldsymbol{r}_0 \times \boldsymbol{E} a d\phi = \int_{-\phi_0}^{\phi_0} h_{zn}'' E_\phi a d\phi, \quad (14b)$$

where $\vec{Y}_n''$ is the admittance seen by the $n$th mode looking in the direction of increasing $r$ from the slot. When (13) and (14) are substituted into (12), the expression (11) for the conductance may be written as

$$G = \frac{\int_{-\phi_0}^{\phi_0} d\phi \int_{-\phi_0}^{\phi_0} d\phi' E_\phi(a, \phi) G''(a; \phi, \phi') E_\phi(a, \phi')}{\left| \int_{-\phi_0}^{\phi_0} h_{z1}'' E_\phi(a, \phi) d\phi \right|^2}, \quad (15)$$

where

$$G''(a; \phi, \phi') = \mathrm{Re} \sum_{n=0}^{\infty} \vec{Y}_n''(a) h_{zn}''(\phi) h_{zn}''(\phi').$$

This expression for the conductance can readily be shown to be stationary with respect to the variation of the aperture field $E_\phi(a, \phi)$ about its correct value.

The expressions for the admittances $\vec{Y}_n(a)$ are given in the Appendix as[7]

$$\vec{Y}_n''(a) = \frac{\kappa}{j\omega\mu a} \frac{H_n^{(2)}(\kappa a)}{H_n^{(2)\prime}(\kappa a)}. \quad (16)$$

Therefore, one finds

$$\mathrm{Re}\, \vec{Y}_n''(a) = \frac{1}{2}\left[\vec{Y}_n''(a) + \vec{Y}_n''^*(a)\right] = \frac{2}{\pi\omega\mu a^2} \frac{1}{\left| H_n^{(2)\prime}(\kappa a) \right|^2}. \quad (17)$$

When the "static" slot electric field,

$$E_\phi = \frac{1}{\sqrt{1 - (\phi/\phi_0)^2}},$$

derived from the integral equation solution presented in the Appendix as (53), is employed as a trial field in (15) and the integration is performed, the unnormalized conductance becomes

$$G = \frac{1}{\pi\omega\mu a^2} \sum_{n=0}^{\infty} \frac{\epsilon_n}{\left| H_n^{(2)\prime}(\kappa a) \right|^2} \left[\frac{J_0(n\phi_0)}{J_0(\phi_0)}\right]^2, \quad (18)$$

where

$$\epsilon_n = 1, \quad n = 0$$
$$= 2, \quad n > 0,$$

and $J_0$ is the Bessel function of the first kind of order zero. When the conductance is normalized in the same manner as was the susceptance in (10), one obtains

$$G' = \frac{G}{Y_0''} = \frac{1}{\pi\kappa a} \sum_{n=0}^{\infty} \frac{\epsilon_n}{\left| H_n^{(2)\prime}(\kappa a) \right|^2} \left[\frac{J_0(n\phi_0)}{J_0(\phi_0)}\right]^2. \quad (19)$$

The complete transverse network for this case is shown in Fig. 3.

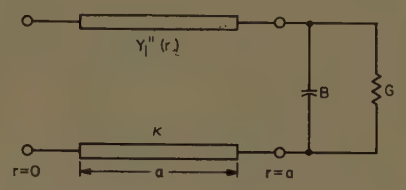



Fig. 3—Transverse equivalent network of slotted circular waveguide for the $n=1$ radial $H$-type mode.

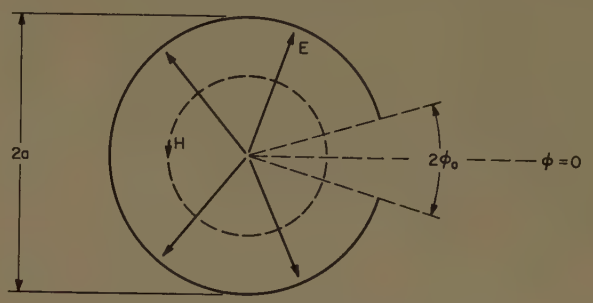



Fig. 4—Slotted circular waveguide carrying the perturbed $E_{01}$ mode.

### B. Perturbed $E_{01}$ Mode

The geometry and magnetic field for this case are shown in Fig. 4. The relevant discontinuity problem is that of the angularly symmetric ($n=0$) $E$-type radial mode incident on the slot from within the waveguide. The determination of the appropriate circuit parameters parallels that in the preceding case. From the results given in the Appendix, one obtains

$$\int_{-\phi_0}^{\phi_0} \boldsymbol{r}_0 \times \boldsymbol{E}_t(a, \phi) a d\phi = j\omega\mu M_\phi I_0' h_0'(a, 0), \quad (20)$$

where

$I_0'$ is the discontinuity in the modal current of the $n=0$ radial $E$-type mode at the slot,

$h_0'(a, 0)$ is the value of the $n=0$ $E$-type mode function at the center of the slot,

$M_\phi = \frac{\pi}{16}\left(\frac{\kappa}{k}\right)^2 (2a\phi_0)^2$ is the magnetic polarizability of the slot for the $\phi$ component of magnetic field,[8] and where the symbols $r_0$ and $E_t$ have the same meaning as in (6).

[7] $H_n^{(2)\prime}$ represents the first derivative of the Hankel function with respect to its argument.

[8] The polarizability in this case becomes identical with that for an infinite slit in an infinite plane, when the width $d$ of the slit in the plane is replaced by the arclength $2a\phi_0$ of the slit in the cylinder. (See C. G. Montgomery, R. H. Dicke, and E. M. Purcell, "Principles of Microwave Circuits," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 8, p. 178, Table 6.1; 1948.)

As before, the mode voltage at the slot can be expressed approximately as

$$V_0' = \boldsymbol{h}_0'^*(a,0)\cdot\int_{-\phi_0}^{\phi_0} \boldsymbol{r}_0 \times \boldsymbol{E}_t(a,\phi)\,ad\phi. \tag{21}$$

The unnormalized susceptance is then given by

$$\frac{I_0'}{V_0'} = -jB = \frac{-j}{\omega\mu M_\phi \boldsymbol{h}_0'(a,0)\cdot\boldsymbol{h}_0'^*(a,0)}$$

$$= \frac{-j8}{\omega\mu\left(\frac{\kappa}{k}\right)^2\phi_0^2}, \tag{22}$$

and the normalized susceptance by

$$B' = \frac{B}{Y_0'} = \frac{8}{\kappa a\phi_0^2}. \tag{23}$$

A variational expression for the conductance in this case may be formulated in a manner similar to that for the $H$-type mode. The resulting expression is

$$G = \frac{\int_{-\phi_0}^{\phi_0} d\phi \int_{-\phi_0}^{\phi_0} d\phi' E_z(a,\phi)G'(a;\phi,\phi')E_z(a,\phi')}{\left|\int_{-\phi_0}^{\phi_0} h_{\phi 0}'E_z(a,\phi)d\phi\right|^2}, \tag{24}$$

where

$$G'(a;\phi,\phi') = \mathrm{Re}\sum_{n=0}^{\infty}\overrightarrow{Y_n}'(a)h_{\phi n}'(a,\phi)h_{\phi n}'(a,\phi').$$

The admittance $\overrightarrow{Y_n}'(a)$ is the dual of the impedance $\overrightarrow{Z_n}''(a)$ given previously and is

$$\overrightarrow{Y_n}'(a) = \frac{j\omega\epsilon a}{\kappa}\frac{H_n^{(2)\prime}(\kappa a)}{H_n^{(2)}(\kappa a)}, \tag{25}$$

so that

$$\mathrm{Re}\,\overrightarrow{Y_n}'(a) = \frac{1}{2}\left[\overrightarrow{Y_n}'(a) + \overrightarrow{Y_n}'^*(a)\right]$$

$$= \frac{2\omega\epsilon}{\pi\kappa^2}\frac{1}{\left|H_n^{(2)}(\kappa a)\right|^2}. \tag{26}$$

When the mode functions from (4) and the "static" trial electric field

$$E_z = \sqrt{1-(\phi/\phi_0)^2},$$

given by (73) of the Appendix are substituted into (24), the integration can be performed to yield the following expression for the unnormalized conductance:

$$G = \frac{8\omega\epsilon}{\pi\kappa^2}\sum_{n=0}^{\infty}\frac{\epsilon_n}{\left|H_n^{(2)}(\kappa a)\right|^2}\left[\frac{J_1(n\phi_0)}{n\phi_0}\right]^2, \tag{27}$$

where

$$\epsilon_n = 1, \quad n = 0$$
$$= 2, \quad n > 0,$$

and $J_1$ is the Bessel function of the first kind of order unity. The normalized conductance is then given by

$$G' = \frac{G}{Y_0'} = \frac{8}{\pi\kappa a}\sum_{n=0}^{\infty}\frac{\epsilon_n}{\left|H_n^{(2)}(\kappa a)\right|^2}\left(\frac{J_1(n\phi_0)}{n\phi_0}\right)^2. \tag{28}$$

The complete transverse network for this case is shown in Fig. 5.

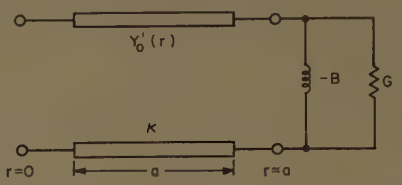



Fig. 5—Transverse equivalent network of slotted circular waveguide for the $n=0$ $E$-type radial mode.

It should be noted that the above expressions for the susceptances given by (9) and (22) are strictly valid only for small-aperture angles ($\sin\phi_0 \approx \phi_0$). However, as shown in the next section, the small-aperture approximation can be pushed somewhat beyond the above restriction. The conductance expressions, while not limited by small-aperture considerations, are practical only when the associated infinite series converge rapidly. The rapidity of convergence depends upon the magnitude of $\kappa a$. For small values of $\kappa a$ the series converge rapidly, but for larger values the convergence is slower. In the two cases treated here the values of $\kappa a$ are of the order of 1.841 and 2.405, respectively. In this range the series converge fairly rapidly; in the calculations of the next section four terms were satisfactory for the $H$-type mode and five terms for the $E$-type mode.

## IV. Perturbation Solutions and Comparison with Measured Values

The propagation constants of the two lowest leaky waves in slotted circular waveguide are obtained from the solutions of the resonance equations for the two transverse equivalent networks derived in the preceding section. The networks are shown in Figs. 3 and 5. The resonance equations for these $H$-type and $E$-type waves are, respectively,

$$\frac{\overleftarrow{Z_1}''(a)}{Z_0''(a)} + R' - jX' = 0 \tag{29}$$

and

$$\frac{\overleftarrow{Z_0}'(a)}{Z_0'(a)} + R' + jX' = 0, \tag{30}$$

where, as shown in the Appendix,

$$\frac{\overleftarrow{Z_1''(a)}}{Z_0''(a)} = -j\frac{J_1'(\kappa a)}{J_1(\kappa a)}$$

and

$$\frac{\overleftarrow{Z_0'(a)}}{Z_0'(a)} = -j\frac{J_0(\kappa a)}{J_0'(\kappa a)}.$$

The normalized values of resistance and reactance $R'$ and $X'$ are those corresponding to the slot discontinuity and are computed from the conductances and susceptances given in (10), (19), (23) and (28) by means of

$$R' = \frac{G'}{G'^2 + B'^2}, \qquad X' = \frac{B'}{G'^2 + B'^2}.$$

It should be recalled that the impedances for the $n=1$ mode are normalized with respect to the $n=0$ mode characteristic impedance for convenience, as discussed in Section III.

As in the case of rectangular leaky waveguides,[1] the resonance equations are complex transcendental equations whose exact solutions require either computing machines or arduous numerical computations. A perturbation technique is therefore employed to obtain simple essentially-closed-form solutions for the propagation constants. A perturbation procedure is also appropriate because most cases of practical interest are concerned with slow leakage.

The unperturbed structure for these cases is taken to be a completely closed circular waveguide. The unperturbed transverse wavenumbers are thus the wavenumbers of the lowest $H$ and $E$ modes in the round guide and are $\kappa_0 a = 1.841$ and $\kappa_0 a = 2.405$, respectively. The perturbed wavenumbers can then be written in the form

$$\kappa = \kappa_0 + \Delta\kappa.$$

When the resonance equations (29) and (30) are expanded about their corresponding unperturbed wavenumbers $\kappa_0$, and the terms containing the derivatives of the terminating impedances are neglected, the perturbation $\Delta\kappa$ for the $H$-type wave is obtained as

$$\Delta\kappa \approx \frac{1}{a}\left(\frac{(\kappa_0 a)^2}{(\kappa_0 a)^2 - 1}\right)(X'(\kappa_0) + jR'(\kappa_0)), \tag{31}$$

with

$$\kappa_0 a = 1.841,$$

and for the $E$-type wave as

$$\Delta\kappa \approx \frac{1}{a}(jR'(\kappa_0) - X'(\kappa_0)), \tag{32}$$

with

$$\kappa_0 a = 2.405.$$

The leaky wave propagation constants are then given by

$$k_z = \beta - j\alpha = \sqrt{k^2 - \kappa^2} \approx \sqrt{k^2 - \kappa_0^2 - 2\kappa_0\Delta\kappa}$$
$$\approx \sqrt{k^2 - \kappa_0^2}\left(1 - \frac{\kappa_0\Delta\kappa}{k^2 - \kappa_0^2}\right).$$

Thus, we may write for the $H$-type wave

$$\frac{\beta}{k} = \frac{\lambda}{\lambda_g}$$
$$\approx \frac{\lambda}{\lambda_{g0}}\left\{1 - \frac{1}{4\pi^2}\left(\frac{\lambda_{g0}}{a}\right)^2\left(\frac{(\kappa_0 a)^3}{(\kappa_0 a)^2 - 1}\right)X'(\kappa_0)\right\}, \tag{33}$$

$$\alpha\lambda \approx \frac{1}{2\pi}\frac{\lambda_{g0}}{\lambda}\left(\frac{\lambda}{a}\right)^2\left(\frac{(\kappa_0 a)^3}{(\kappa_0 a)^2 - 1}\right)R'(\kappa_0), \tag{34}$$

with

$$\lambda_{g0} = \frac{2\pi}{\sqrt{k^2 - \kappa_0^2}},$$

and for the $E$-type wave

$$\frac{\beta}{k} = \frac{\lambda}{\lambda_g} \approx \frac{\lambda}{\lambda_{g0}}\left\{1 + \frac{\kappa_0 a}{4\pi^2}\left(\frac{\lambda_{g0}}{a}\right)^2 X'(\kappa_0)\right\}, \tag{35}$$

$$\alpha\lambda \approx \frac{\kappa_0 a}{2\pi}\frac{\lambda_{g0}}{\lambda}\left(\frac{\lambda}{a}\right)^2 R'(\kappa_0), \tag{36}$$

with

$$\lambda_{g0} = \frac{2\pi}{\sqrt{k^2 - \kappa_0^2}}.$$

In Figs. 6–9 theoretical curves computed from (33)–(36) are compared with measurements published by Harrington.[2] Since the susceptance expressions employed in the resonance relations are based on small-aperture solutions, it is to be expected that the smaller angle results in Figs. 6–9 agree well with the measured values, while those for the larger angles show a discrepancy.

Harrington's theoretical results[2] are not shown here, but his analysis is similar to Rumsey's work on rectangular waveguides[9] in which a variational procedure is used that is equivalent to the determination of the equivalent circuit parameters of the slot. The trial fields used by Harrington differ from those used in this paper for the evaluation of the radiation conductances in that the fields he uses do not satisfy appropriate edge conditions. In calculating the propagation constants, Harrington has also employed a perturbation technique, but he retains terms of the order of $(\Delta\kappa)^2$. Consequently, his results for the attenuation constant in the case of wide slots are superior to those computed from (34) and (36), but at the cost of a substantial increase in the complexity and difficulty of actual computation such that no explicit expressions for the attenuation constants can be obtained.

[9] V. H. Rumsey, "Traveling wave slot antennas," *J. Appl. Phys.*, vol. 24, pp. 1358–1365; November, 1953.

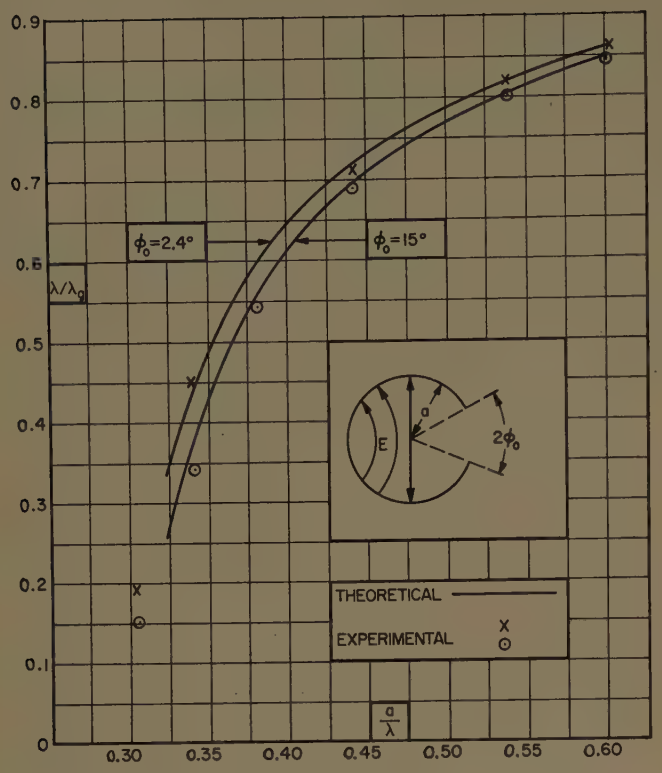



Fig. 6—Guide wavelength vs guide radius for *H*-type leaky wave in circular guide.

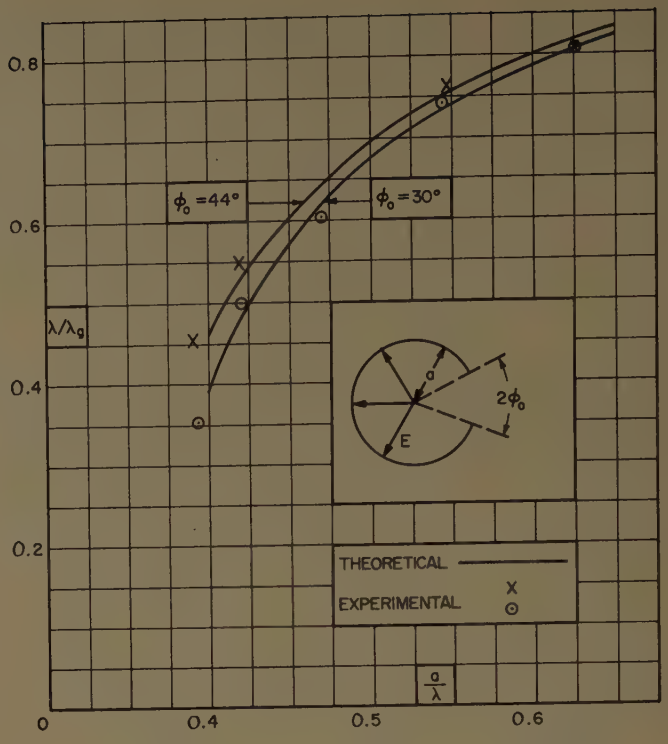



Fig. 8—Guide wavelength vs guide radius for *E*-type leaky wave in circular guide.

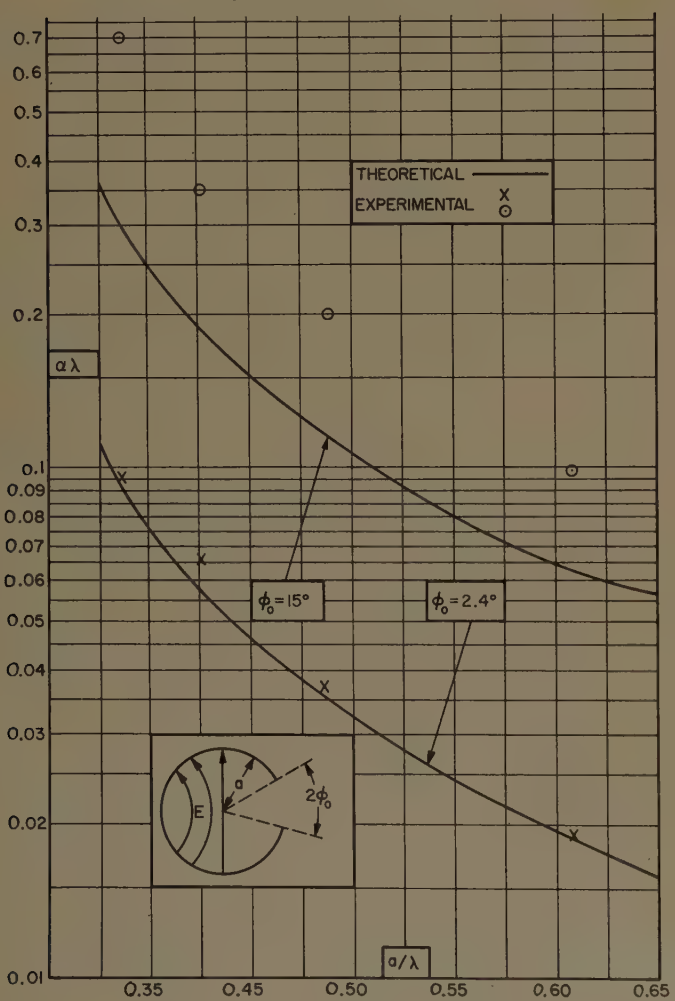



Fig. 7—Attenuation constant vs guide radius for *H*-type leaky wave in circular guide.

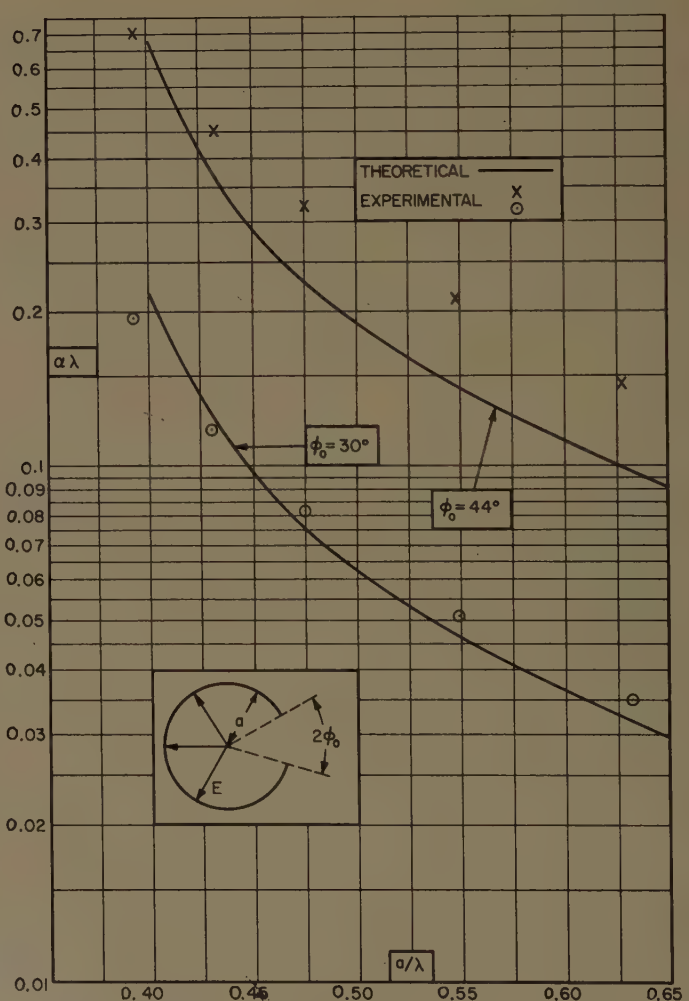



Fig. 9—Attenuation constant vs guide radius for *E*-type leaky wave in circular guide.

## Appendix

## Diffraction of a Radial Wave by a Slotted Circular Waveguide

The equivalent circuit representations of the slotted circular waveguide, which are employed in the transverse resonance calculation for the leaky wave propagation constants, require the solutions of two diffraction problems, one for each type of leaky wave. These two correspond to the $n=1$ $H$-type radial mode and $n=0$ $E$-type radial mode incident on the slot from within the circular waveguide.

Each of the two diffraction problems is stated in terms of an integral equation for the electric field in the slot aperture. The integral equations are solved in the static and small-aperture limits, and expressions are derived for the aperture fields and the slot polarizabilities.

### *A. H-Type Radial Mode Incident* ($n=1$)

The radial $H$-type mode incident from within the circular waveguide possesses an angular dependence $\cos\phi$ and an axial dependence $\exp(-jk_z z)$. The transverse (to $r$) fields of this incident mode are given by[10]

$$\begin{aligned} \boldsymbol{E}_{t\text{inc}}(r,\phi) &= V_{\text{inc}}''(r)\boldsymbol{e}_1''(r,\phi) \\ \boldsymbol{H}_{t\text{inc}}(r,\phi) &= I_{\text{inc}}''(r)\boldsymbol{h}_1''(r,\phi), \end{aligned} \tag{37}$$

where the mode functions $\boldsymbol{e}_1''$ and $\boldsymbol{h}_1''$ are those given in (5) for $n=1$. If the region within the guide ($r<a$) is designated as region 1 and that outside ($r>a$) as region 2, the transverse (to $r$) magnetic fields in the two regions can be represented in terms of the complete set of $H$-type vector mode functions as

$$\begin{aligned} \boldsymbol{H}_t^{(1)}(r,\phi) &= \sum_{n=0}^{\infty} I_n^{(1)''}(r)\boldsymbol{h}_n''(r,\phi) \\ \boldsymbol{H}_t^{(2)}(r,\phi) &= \sum_{n=0}^{\infty} I_n^{(2)''}(r)\boldsymbol{h}_n''(r,\phi). \end{aligned} \tag{38}$$

The imposition of the continuity requirement on the tangential magnetic field at the aperture then leads to an integral equation for the transverse electric field in the aperture as follows: Continuity of the tangential magnetic field at the aperture requires that

$$\boldsymbol{H}_t^{(1)}(a,\phi) = \boldsymbol{H}_t^{(2)}(a,\phi).$$

When these fields are expanded in terms of the radial modes,

$$\sum_{n=0}^{\infty} I_n^{(1)''}(a)\boldsymbol{h}_n(a,\phi) = \sum_{n=0}^{\infty} I_n^{(2)''}(a)\boldsymbol{h}_n(a,\phi)$$

is obtained, where $I_n^{(1)''}(a)$ and $I_n^{(2)''}(a)$ are the modal currents at $r=a^-$ and $r=a^+$, respectively. This expression can now be rearranged slightly to yield

$$I_1''(a)\boldsymbol{h}_1''(a,\phi) = \sum_{n=0}{}' I_n''(a)\boldsymbol{h}_n(a,\phi), \tag{39}$$

where

$$\begin{aligned} I_1''(a) &= I_1^{(1)''}(a) - I_1^{(2)''}(a) \\ I_n''(a) &= I_n^{(1)''}(a) - I_n^{(2)''}(a), \end{aligned}$$

and the prime on the summation indicates the omission of the $n=1$ term.

For all except the $n=1$ mode,

$$\begin{aligned} I_n^{(2)''}(a) &= \overrightarrow{Y}_n''(a)V_n''(a) \\ I_n^{(1)''}(a) &= -\overleftarrow{Y}_n''(a)V_n''(a). \end{aligned} \tag{40}$$

The admittances $\overrightarrow{Y}_n''$ and $\overleftarrow{Y}_n''$ are the admittances "seen" by the $n$th $H$-type mode at $r=a$, looking in the direction of increasing $r$ and decreasing $r$, respectively. These two admittances can be shown to be

$$\begin{aligned} \overrightarrow{Y}_n''(a) &= \frac{\kappa}{j\omega\mu a}\frac{H_n^{(2)}(\kappa a)}{H_n^{(2)\prime}(\kappa a)}, \\ \overleftarrow{Y}_n''(a) &= \frac{\kappa}{j\omega\mu a}\frac{J_n(\kappa a)}{J_n'(\kappa a)}, \end{aligned}$$

where $\kappa=\sqrt{k^2-k_z^2}$. It should be noted that these are not the radial line characteristic admittances since on a radial transmission line, unlike a uniform line, the actual current-voltage ratio is not equal to the characteristic admittance even under matched conditions.

The modal voltages $V_n''(a)$ are given by the transforms of the transverse (to $r$) electric field as

$$\begin{aligned} V_n''(a) &= \int_{-\phi_0}^{\phi_0} \boldsymbol{h}_n''^{*} \times \boldsymbol{r}_0 \cdot \boldsymbol{E}_t(a,\phi')a\,d\phi' \\ &= \int_{-\phi_0}^{\phi_0} h_{zn}''(\phi')E_\phi(a,\phi')a\,d\phi'. \end{aligned} \tag{41}$$

By combining (39), (40), and (41), one obtains the integral equation

$$\begin{aligned} &I_1''(a)h_{z1}''(a,\phi) \\ &= \int_{-\phi_0}^{\phi_0} \sum_{n=0}{}' \overleftrightarrow{Y}_n''(a)h_{zn}''(a,\phi)h_{zn}''(a,\phi')E_\phi(a,\phi')a\,d\phi', \end{aligned} \tag{42}$$

where

$$\overleftrightarrow{Y}_n'' = \overrightarrow{Y}_n'' + \overleftarrow{Y}_n''.$$

The integral equation (42) may be rewritten as

$$\begin{aligned} &[I_1''(a) + \overleftrightarrow{Y}_1''(a)V_1''(a)]h_{z1}''(a,\phi) - \overleftrightarrow{Y}_0''(a)V_0''(a)h_{z0}''(a,\phi) \\ &= \int_{-\phi_0}^{\phi_0} \sum_{n=1}^{\infty} \overleftrightarrow{Y}_n''(a)h_{zn}''(a,\phi)h_{zn}''(a,\phi')E_\phi(a,\phi')a\,d\phi'. \end{aligned} \tag{43}$$

[10] The exponential $z$ dependence is the same for all field components and is suppressed.

In order to recast the above integral equation into a form permitting diagonalization of the kernel a set of "static" admittances are introduced as follows:

$$\overleftrightarrow{Y}_{ns}''(a) = -\frac{2\kappa^2}{j\omega\mu n},$$

so that

$$\overleftrightarrow{Y}_{n}''(a) - \overleftrightarrow{Y}_{ns}''(a) \to 0 \qquad \text{for } n \gg \kappa a.$$

The integral equation (43) can also be rewritten, employing a compact notation patterned after that used in the "Waveguide Handbook,"[11] in the form

$$\sum_{m=0}^{\infty} \hat{I}_m h_{zm}''(a, \phi) = \int_{-\phi_0}^{\phi_0} \sum_{n=1}^{\infty} \overleftrightarrow{Y}_{ns}''(a) h_{zn}''(a, \phi) h_{zn}''(a, \phi') E_\phi(a, \phi') a d\phi', \quad (44)$$

where

$$\hat{I}_m = -\overleftrightarrow{Y}_0'' V_0'', \qquad m = 0,$$
$$\hat{I}_m = I_1'' + \overleftrightarrow{Y}_1'' V_1'', \qquad m = 1,$$
$$\hat{I}_m = (\overleftrightarrow{Y}_{ms}'' - \overleftrightarrow{Y}_{m}'') V_m'', \qquad m > 1.$$

In view of the linearity of (44) the electric field in the aperture can be represented in terms of partial fields as

$$E_\phi(a, \phi) = \sum_{m=0}^{\infty} \hat{I}_m(a) \mathcal{E}_{\phi m}(\phi).$$

When the slot aperture is narrow ($\sin \phi_0 \approx \phi_0$), the set of integral equations (44) can be simplified by approximating the effective exciting currents $\hat{I}_m$ in the following manner:

$$\hat{I}_1 = I_1'', \qquad m = 1$$
$$\hat{I}_m = 0, \qquad m \neq 1.$$

The justification for this approximation is the following. It is to be noted that the true $\overleftrightarrow{Y}_n$ values are complex; in fact, the few lowest ones, which contribute significantly to the far-field radiation pattern, are predominantly real in character. Alternatively stated, the radiated power is not carried by the $n=1$ mode alone. As a result, the admittance derived here for the slot actually should contain a conductive contribution. The susceptance value of a narrow slit, on the other hand, consists largely of contributions from the high modes, which are essentially purely reactive and for which $Y_n = Y_{ns}$ is an excellent approximation. Furthermore, while it is true that for the lowest modes the difference between $Y_n$ and $Y_{ns}$ is largest, these modes do not have much reactive content and will not contribute significantly to the susceptance of a narrow slit. For these reasons, the procedure below will result in a purely susceptive slot admittance which can be expected to be a reasonably accurate approximation to the slot susceptance. The radiation conductance value obtained on this basis, however, would be in serious error. A variational procedure is therefore employed, in Section III, to evaluate the radiation conductance values in a separate and accurate fashion.

With the approximations indicated above, the infinite set of integral equations (44) reduces to the single integral equation

$$h_{z1}(a, \phi) = \int_{-\phi_0}^{\phi_0} \sum_{n=1}^{\infty} \overleftrightarrow{Y}_{ns}''(a) h_{zn}''(a, \phi) h_{zn}''(a, \phi') \mathcal{E}_{\phi 1}(\phi') a d\phi'. \quad (45)$$

When the mode functions and static admittances are inserted into (45), the integrand is seen to be an even function of $\phi$, and the equation assumes the form

$$\frac{\cos \phi}{\sqrt{\pi}} = -\frac{4\kappa^2 a}{j\omega\mu\pi} \int_0^{\phi_0} \sum_{n=1}^{\infty} \frac{\cos n\phi \cos n\phi'}{n} \mathcal{E}_{\phi 1}(\phi') d\phi'. \quad (46)$$

In order to diagonalize the kernel of (46), the following identity[12] is invoked:

$$-2 \sum_{n=1}^{\infty} \frac{\cos n\phi \cos n\phi'}{n} = \ln 2 \left| \cos \phi - \cos \phi' \right|, \quad (47)$$

and a change of variable is made according to

$$\cos \phi = \alpha \cos \theta + \beta.$$

Since $0 < \phi < \phi_0$, we have that $0 < \theta < \pi$, and

$$\alpha = \sin^2 \phi_0/2, \qquad \beta = \cos^2 \phi_0/2. \quad (48)$$

Eq. (46) may then be written as

$$\alpha \cos \theta + \beta = \frac{2\kappa^2}{j\omega\mu} \frac{a}{\sqrt{\pi}} \int_0^{\pi} \ln 2\alpha \left| \cos\theta - \cos\theta' \right| \hat{\mathcal{E}}_{\phi 1}(\theta') d\theta', \quad (49)$$

where

$$\hat{\mathcal{E}}_{\phi 1}(\theta') = \mathcal{E}_{\phi 1}(\theta') \frac{d\theta'}{d\phi'}.$$

Since the aperture is assumed to be small,

$$\alpha \approx \phi_0^2/4, \qquad \beta \approx 1.$$

With the above approximations and the use of the identity, (48) and (49) may finally be written as

$$1 \approx \frac{2\kappa^2}{j\omega\mu} \frac{a}{\sqrt{\pi}} \int_0^{\pi} \left[ \ln \alpha - 2 \sum_{n=1}^{\infty} \frac{\cos n\theta \cos n\theta'}{n} \right] \hat{\mathcal{E}}_{\phi 1}(\theta') d\theta'. \quad (50)$$

[11] *Op. cit.*, Marcuvitz, "Waveguide Handbook," Section 3.5.

[12] *Cf.*, *ibid.*, p. 147.

If $\hat{\mathcal{E}}_{\phi 1}$ is now represented as a Fourier cosine series,

$$\hat{\mathcal{E}}_{\phi 1}(\theta') = \sum_{n=0}^{\infty} a_n \cos n\theta', \tag{51}$$

it follows immediately from (50) that

$$a_0 = -j\left(\frac{k}{\kappa}\right)^2 \frac{1}{4\omega\epsilon a\sqrt{\pi}\ln(2/\phi_0)} \tag{52}$$

and

$$a_n = 0, \qquad n > 0.$$

It follows from (51) and (52) that for a narrow slot ($\sin\phi_0 \approx \phi_0$) the electric field in the aperture is given approximately by

$$E_\phi(\phi) \approx \frac{1}{j\omega\epsilon}\left(\frac{k}{\kappa}\right)^2 \frac{I_1''}{2a\sqrt{\pi}} \frac{1}{\ln(2/\phi_0)} \frac{1}{\sqrt{\phi_0^2 - \phi^2}}. \tag{53}$$

It is also of interest to note that the solution for the field satisfies the edge condition.

The magnetic polarizability of the slot for this case can be defined as follows:

$$\int_{-\phi_0}^{\phi_0} \boldsymbol{r}_0 \times \boldsymbol{E}_t(\phi) a d\phi = \boldsymbol{z}_0 \int_{-\phi_0}^{\phi_0} E_\phi(\phi) a d\phi$$

$$= \boldsymbol{z}_0 \frac{\sqrt{\pi}}{j\omega\epsilon}\left(\frac{k}{\kappa}\right)^2 \frac{I_1''}{2\ln(2/\phi_0)}. \tag{54}$$

Eq. (54) can then be written in the following form:

$$\int_{-\phi_0}^{\phi_0} \boldsymbol{r}_0 \times \boldsymbol{E}_t(\phi) a d\phi = \frac{1}{j\omega\epsilon} M_z I_1'' \boldsymbol{h}_1''(a, 0), \tag{55}$$

where

$$\boldsymbol{h}_1''(a, 0) = \boldsymbol{z}_0\, 1/\sqrt{\pi};$$

upon comparison with (54) we find that

$$M_z = \frac{\pi}{2}\left(\frac{k}{\kappa}\right)^2 \frac{1}{\ln(2/\phi_0)} \tag{56}$$

is the axial magnetic polarizability of the slot in the cylinder.

### B. E-Type Radial Mode Incident ($n=0$)

In this case the transverse (to $r$) fields of the incident mode are given by

$$\boldsymbol{E}_{t\,\mathrm{inc}}(r, \phi) = V_{\mathrm{inc}}'(r)\boldsymbol{e}_0'(r, \phi),$$
$$\boldsymbol{H}_{t\,\mathrm{inc}}(r, \phi) = I_{\mathrm{inc}}'(r)\boldsymbol{h}_0'(r, \phi), \tag{57}$$

where the mode functions $\boldsymbol{e}_0'$ and $\boldsymbol{h}_0'$ are given in (4) with $n=0$.

When the external and internal tangential magnetic fields at the slit are equated as in the preceding case the following integral equation is obtained:

$$I_0' h_{\phi 0}'(a, \phi)$$

$$= -\int_{-\phi_0}^{\phi_0} \sum_{n=1}^{\infty} \overleftrightarrow{Y}_n'(a) h_{\phi n}'(a, \phi) h_{\phi n}'(a, \phi') E_z(\phi') a d\phi', \tag{58}$$

where

$$I_0' = I_0^{(1)'} - I_0^{(2)'}$$

is the discontinuity in the current of the $n=0$ mode at $r=a$, and

$$\overleftrightarrow{Y}_n'(a) = \overleftarrow{Y}_n'(a) + \overrightarrow{Y}_n'(a)$$

is the dual of that in the first section of this Appendix. The admittances $\overrightarrow{Y}_n'(a)$ and $\overleftarrow{Y}_n'(a)$ are therefore given by

$$\overrightarrow{Y}_n'(a) = \frac{j\omega\epsilon a}{\kappa} \frac{H_n^{(2)'}(\kappa a)}{H_n^{(2)}(\kappa a)},$$
$$\overleftarrow{Y}_n'(a) = -\frac{j\omega\epsilon a}{\kappa} \frac{J_n'(\kappa a)}{J_n(\kappa a)}. \tag{59}$$

The "static" admittances for this case are

$$\overrightarrow{Y}_{ns}'(a) = \overleftarrow{Y}_{ns}'(a) = -\frac{j\omega\epsilon n}{\kappa^2}. \tag{60}$$

Integral equation (58) can then be written as

$$\sum_{m=0}^{\infty} \hat{I}_m h_{\phi m}'(a, \phi)$$

$$= -\int_{-\phi_0}^{\phi_0} \sum_{n=1}^{\infty} \overleftrightarrow{Y}_{ns}'(a) h_{\phi n}'(a, \phi) h_{\phi n}'(a, \phi') E_z(\phi') a d\phi', \tag{61}$$

where

$$\hat{I}_m = I_0', \qquad m = 0,$$
$$\hat{I}_m = (\overleftrightarrow{Y}_m'(a) - \overleftrightarrow{Y}_{ms}'(a)) V_m', \qquad m > 0.$$

As before, the aperture field can be represented in terms of a set of partial fields,

$$E_z(\phi) = -\sum_{m=0}^{\infty} \hat{I}_m \mathcal{E}_{zm}(\phi). \tag{62}$$

If higher exciting currents $\hat{I}_m$ are again neglected in comparison with $\hat{I}_0 = I_0'$, the set of integral equations reduces to

$$1 = j\sqrt{\frac{8}{\pi}} \frac{\omega\epsilon}{\kappa^2} \int_{-\phi_0}^{\phi_0} \sum_{n=0'}^{\infty} n \cos n\phi \cos n\phi' \mathcal{E}_{z0}(\phi') d\phi'. \tag{63}$$

The integral equation (63) may now be solved in the following manner: First, we recognize that

$$\int_{-\phi_0}^{\phi_0} \sum_{n=1}^{\infty} n \cos n\phi \cos n\phi' \mathcal{E}_{z0}(\phi') d\phi'$$

$$= 2\int_0^{\phi_0} \sum_{n=1}^{\infty} n \cos n\phi \cos n\phi' \mathcal{E}_{z0}(\phi') d\phi'.$$

If the above integral is now integrated by parts, we have

$$2\int_0^{\phi_0}\sum_{n=1}^{\infty} n\cos n\phi\cos n\phi'\mathcal{E}_{z0}(\phi')d\phi' = 2\sum_{n=1}^{\infty}\cos n\phi\sin n\phi'\mathcal{E}_{z0}(\phi')\Big]_0^{\phi_0} + 2\int_0^{\phi_0}\frac{d}{d\phi'}\sum_{n=1}^{\infty}\frac{\cos n\phi\cos n\phi'}{n}\frac{d}{d\phi'}\mathcal{E}_{z0}(\phi')d\phi'.$$

Since $\mathcal{E}_{z0}(\phi)$ vanishes at $\phi=\phi_0$,

$$\int_{-\phi_0}^{\phi_0}\sum_{n=1}^{\infty} n\cos n\phi\cos n\phi'\mathcal{E}_{z0}(\phi')d\phi' = 2\int_0^{\phi_0}\frac{d}{d\phi'}\sum_{n=1}^{\infty}\frac{\cos n\phi\cos n\phi'}{n}\frac{d}{d\phi'}\mathcal{E}_{z0}(\phi')d\phi'.$$

Employing the identity (47), we see that

$$\frac{d}{d\phi'}\sum_{n=1}^{\infty}\frac{\cos n\phi\cos n\phi'}{n} = -\frac{1}{2}\frac{\sin\phi'}{\cos\phi-\cos\phi'}.$$

Eq. (63) can now be written as[13]

$$1 = -j\sqrt{\frac{8}{\pi}}\frac{\omega\epsilon}{\kappa^2}\int_0^{\phi_0}\frac{\sin\phi'}{\cos\phi-\cos\phi'}\frac{d}{d\phi'}\mathcal{E}_{z0}(\phi')d\phi'. \quad (64)$$

The following change of variable is now made:

$$\cos\phi = \alpha\cos\theta+\beta,$$

where, for $0\le\phi\le\phi_0$, we have, as before, $0\le\theta\le\pi$ with

$$\beta = \cos^2\phi_0/2, \qquad \alpha = \sin^2\phi_0/2,$$

so that

$$1 = \int_0^{\pi}\frac{g(\theta')}{\cos\theta-\cos\theta'}d\theta', \quad (65)$$

where

$$g(\theta') = -j\sqrt{\frac{8}{\pi}}\frac{\omega\epsilon}{\kappa^2}\sin\theta'\frac{d}{d\phi'}\mathcal{E}_{z0}(\phi'). \quad (66)$$

The function $g(\theta')$ may be represented as

$$g(\theta') = \sum_{\nu=0}^{\infty} a_\nu\cos\nu\theta'. \quad (67)$$

Then, since[14]

$$\int_0^{\pi}\frac{\cos\nu\theta'}{\cos\theta-\cos\theta'}d\theta' = -\pi\frac{\sin\nu\theta}{\sin\theta}, \quad (68)$$

(65) becomes

$$1 = -\pi\sum_{\nu=1}^{\infty} a_\nu\frac{\sin\nu\theta}{\sin\theta}. \quad (69)$$

[13] This integral exists in the meaning of Cauchy's principal value.
[14] W. Magnus and F. Oberhettinger, "Special Functions of Mathematical Physics," Chelsea Publishing Co., New York, N. Y., p. 141; 1949.

Since, from (66),

$$g(0) = 0,$$

we note from (67) that

$$\sum_{\nu=0}^{\infty} a_\nu = a_0 + \sum_{\nu=0}^{\infty} a_\nu = 0. \quad (70)$$

From (69) and (70), therefore,

$$a_0 = -a_1 = 1/\pi, \qquad a_\nu = 0,\ \nu > 1,$$

and as a result, we find

$$g(\theta) = -j\sqrt{\frac{8}{\pi}}\frac{\omega\epsilon}{\kappa^2}\frac{d\mathcal{E}_{z0}}{d\phi}\sin\theta = \frac{1}{\pi}(1-\cos\theta). \quad (71)$$

Since the aperture is assumed to be narrow, $\sin\phi_0\approx\phi_0$, it follows from (66) and (71) that

$$\frac{d\mathcal{E}_{z0}}{d\phi}\approx j\frac{\kappa^2}{\omega\epsilon}\frac{1}{\sqrt{8\pi}}\frac{\phi}{\sqrt{\phi_0^2-\phi^2}}. \quad (72)$$

When the differential equation (72) is integrated, the result is

$$\mathcal{E}_{z0}(\phi) = -j\frac{\kappa^2}{\omega\epsilon}\frac{1}{\sqrt{8\pi}}\sqrt{\phi_0^2-\phi^2}+K;$$

however, since $\mathcal{E}_{z0}(\phi_0)=0$, $K=0$. In view of (62), the aperture electric field is therefore given by

$$E_z(\phi) = j\frac{\kappa^2}{\omega\epsilon}\frac{I_0'}{\sqrt{8\pi}}\sqrt{\phi_0^2-\phi^2}. \quad (73)$$

The magnetic polarizability of the slot for this case is then obtained by writing

$$\int_{-\phi_0}^{\phi_0} r_0\times E_i(\phi)ad\phi = -\phi_0\int_{-\phi_0}^{\phi_0}E_z(\phi)ad\phi = -\phi_0\frac{j\omega\mu}{\sqrt{8\pi}}\left(\frac{\kappa}{k}\right)^2 I_0'\left(\frac{\pi a\phi_0^2}{2}\right) = j\omega\mu M_\phi I_0' h_0'(a,0),$$

where

$$h_0'(a,0) = -\phi_0 1/a\sqrt{2\pi};$$

as a result, we have that

$$M_\phi = \frac{\pi}{16}\left(\frac{\kappa}{k}\right)^2(2a\phi_0)^2 \quad (74)$$

is the magnetic polarizability for magnetic field excitation transverse to the slot.


## Acknowledgment

The authors wish to express their thanks to Mrs. A. Lotito for performing the numerical computations.

# A Spiral-Grating Array*

J. R. DONNELLAN† AND R. T. CLOSE†

*Summary*—**A problem associated with a two-dimensional spiral doublet array is concerned with the fact that use of the spiral doublet as an element to obtain linear polarization leads to spacing difficulties, since the distance between elements is always twice as much in one dimension as in the other. Thus, the allowable scan is seriously limited in one principal plane. A method is proposed for obtaining linear polarization from a *single* spiral antenna by using it in combination with a polarization grating and a ground plane. Experimental results were obtained by combining eight of these elements in a one-dimensional array. For the broadside position of the main beam, the field was found to be linear, with the cross-polarized component of the main beam being 40 db down from the principal polarization. As the beam was scanned out to ±45°, the linearity decreased quite rapidly (to approximately 10 db at 45°) and the plane of the linear polarization was rotated; sidelobes were higher than predicted. Use of this array rather than an array of doublets would involve a compromise concerning specifications, dependent upon the particular application of the array.**

## Introduction

A SUCCESSFUL scanning array employing flat two-wire Archimedean spirals has already been developed at the U. S. Naval Research Laboratory.[1] In order to obtain a linearly-polarized beam, this system relied upon the properties of the spiral doublet: two equally excited spirals of opposite sense placed side by side. However, use of the doublet as an element leads to a spacing problem in a two-dimensional array: the spacing between elements is always twice as much in one dimension as in the other. Thus, the maximum attainable angle of scan (*i.e.*, the maximum angle to which the beam can be scanned without the introduction of an extraneous main lobe) is seriously limited in one principal plane.

Several proposals that appear to remedy this situation have been considered. For example, the use of the doublet configuration can be retained if the size of the individual spiral elements can be reduced, or if, by some arrangement of elements, the spacing does not become excessive. The system described in this report is the result of an attempt to eliminate the need for the doublet in order to produce linear polarization from an array of spirals. It makes use of a polarization grating of parallel wires[2] to produce the requisite linear polarization.

* Received by the PGAP, July 15, 1960; revised manuscript received, October 12, 1960. The research was supported by the AF Cambridge Res. Ctr., Air Res. and Dev. Command, under Contract MIPR-59-714.

† U. S. Naval Res. Lab., Washington, D. C.

[1] J. R. Donnellan, "An Eight Spiral Doublet Electromechanical Scanning Array," Naval Res. Lab., Washington, D. C., Rept. No. 5283; April 10, 1959.

[2] J. R. Donnellan and B. Sheleg, invention disclosure entitled, "A Single Spiral Linearly Polarized Antenna," Naval Res. Lab., Washington, D. C., Docket No. NRL 3638; June 4, 1959.

## Theoretical Considerations

The following discussion demonstrates a method of obtaining linear polarization from a *single* spiral antenna by using it in combination with a polarization grating and a ground plane. The grating consists of closely spaced wires stretched on a frame, the spacing between the wires being much less than a wavelength and much greater than the wire diameter. Consider now a vertical grating situated a quarter-wavelength, at the frequency of operation, above a ground plane and a spiral situated at an arbitrary distance $d$ in front of the grating (Fig. 1). When the spiral is properly excited at its terminals, it will radiate equal amounts of circularly-polarized radiation (of opposite senses) in both the forward and backward directions. The sense of circular polarization is that observed when the radiation is viewed in the direction of propagation.

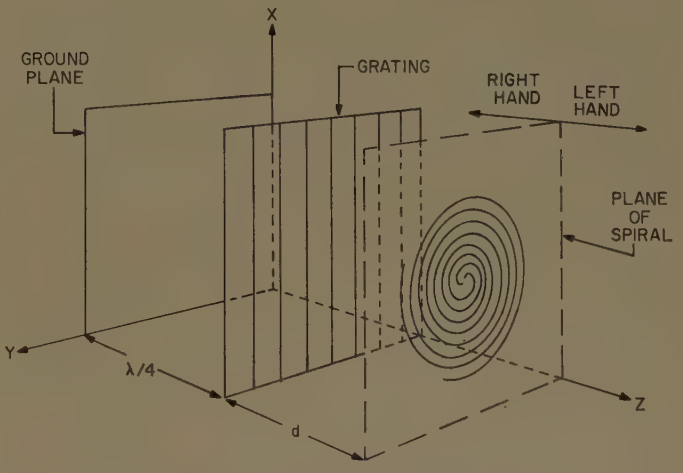



Fig. 1—Spiral-grating configuration.

Assume, for example, that the spiral radiates left-hand circular polarization in the $+Z$ direction and right-hand circular polarization in the $-Z$ direction. Consider the radiation (right-hand) which travels toward the grating. The circular polarization can be regarded as the vector sum of horizontal and vertical components in phase quadrature. The vertical (parallel to the $X$ axis) component travels a distance $d$, is reflected at the grating with a 180° phase change, and returns to the spiral, again traveling a distance $d$, so that the "effective path length" traveled is $2d+\lambda/2$. The horizontal (parallel to the $Y$ axis) component passes through the grating after traveling a distance $d$, then travels a distance $\lambda/4$, is reflected from the ground

plane with a 180° phase change, again travels a distance $\lambda/4$, and finally a distance $d$, returning to the spiral; the "effective path length" traveled is $2d+\lambda$. Thus, the phase angle between the horizontal and vertical components has changed from $+90°$ to $-90°$; *i.e.*, the combination of grating and ground plane serves to retain the rotational sense of the radiation incident upon it. This reflected radiation of one sense will combine with the direct radiation, which is of the other sense, to produce linear polarization. The orientation of the plane of linear polarization is a function of $d$ for a fixed grating position. For a vertical grating, if $d=\lambda/4$, vertical polarization will be obtained (if $d=\lambda/8$, 45° polarization is obtained, etc.); for a horizontal grating (with $d=\lambda/4$), horizontal polarization is obtained. It can also be shown that, with the spiral fixed at a distance $\lambda/2$ above the ground plane, the spacing between grating and ground plane can be varied to produce linear polarization on-axis and at one other angle off-axis (this was actually done in the experimental work).

Thus, with the grating in a fixed position and the spiral a fixed distance above the grating, the spiral may be rotated in order to change the phase of its far field while maintaining linear polarization. That is, rotation of the spiral through an angle $\theta$ advances the phase of both the right-hand and left-hand circular polarization by an amount $\theta$; therefore, the plane of the linear polarization is maintained, while the phase of the far field is changed by an amount $\theta$. A scanning array could thus be composed of many of these elements, and equal spacing between elements would be maintained in both dimensions.

## Experimental Considerations

Preliminary experiments were carried out with a single spiral and grating in which the spacings between spiral and grid, and between grid and ground plane, could be varied. These experiments verified that this system would indeed provide linear polarization, while maintaining the phasing properties of the spiral. It was found that, off axis, the plane of linear polarization rotated from that of the broadside position. This was due to the fact that the effective spacing between grid and ground plane for an off-axis position is greater than the spacing for the broadside position; thus an additional relative phase change between the right-hand and left-hand circular polarizations is produced and the plane of linear polarization is rotated. However, it was decided that an array would nevertheless be built since rotation of the plane of polarization might not be a serious disadvantage, depending on the particular application. It was experimentally determined that, in order to obtain suitable linearity for both the broadside and off-axis positions (out to approximately $\pm 45°$), the spacing between spiral and grating should be approximately $\lambda/8$ and the spacing between grating and ground plane, approximately $3\lambda/8$.

A photograph of the eight-spiral array, operated at a frequency of 1420 Mc, is shown in Fig. 2. The array consisted of one row of spirals (all of the same sense) placed in front of a horizontal grid structure. The feedboard was a microstrip ring network type, in which the desired aperture illumination was obtained by varying the lengths of line between the ring networks. A diagram of this feedboard is shown in Fig. 3. It was designed to give a 25-db Tchebycheff illumination working into matched loads. The grating was constructed of no. 24 (1/32-inch diameter) tinned copper wire with a $\frac{1}{2}$-inch spacing between wires. The lateral spacing between spirals was $\lambda/2$.

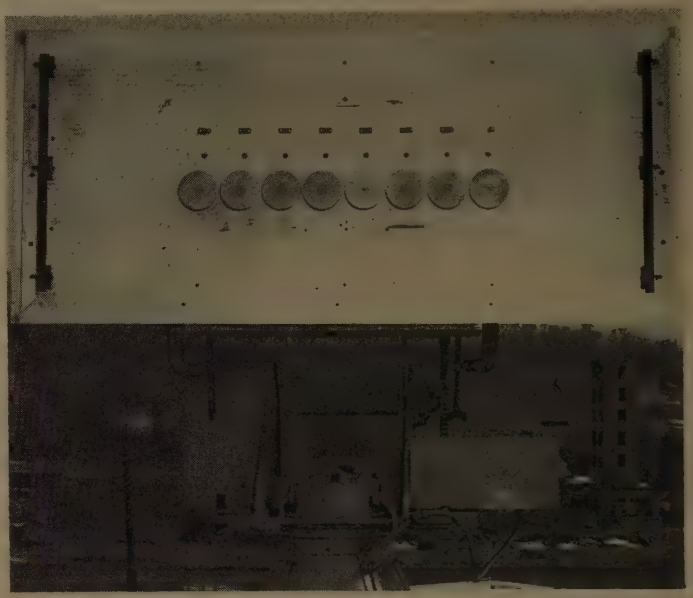


Fig. 2—Spiral-grating array.

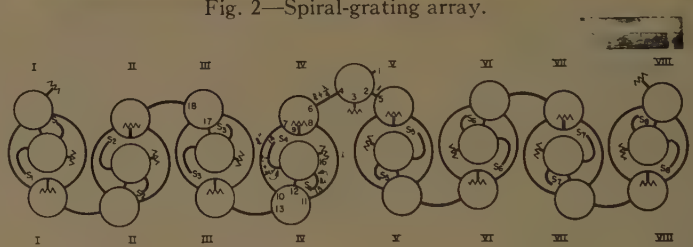


Fig. 3—Microstrip ring network feedboard.

An additional feature of this array was the use of "plug-in" spirals; *i.e.*, two small male conductors (UG-260-U) were attached to the input terminals of each spiral, while two small female connectors (UG-260-U) were attached to each two-wire line. The spirals could then be easily removed from, or inserted in, the array by simply pulling them out or plugging them back in. In the eight-doublet array of Donnellan,[1] the spirals were soldered to the two-wire lines, which made removal or replacement of the spirals a rather lengthy task.

The array was phased for the broadside position by using it as the receiver, while a pyramidal horn was used as the transmitter. The distance between transmitter and receiver was approximately 120 feet (about $5(2D^2/\lambda)$

at 1420 Mc) and the horn could be rotated to provide any arbitrary linear polarization. Two elements at a time were connected into the feed system (matched loads terminating all other connections) and one spiral was rotated with respect to the other until both were in phase (the technique is described more completely in Donnellan.[1] This process was continued for all elements taken two at a time until all were in phase. A second row of elements was actually inserted several times and measurements were taken. This series of measurements indicated that two-dimensional arrays without a spacing problem were feasible (with certain limitations, as will be seen from the following antenna patterns). Patterns of single and double rows were found to be essentially the same, so that, in the interest of ease in experimentation, pattern measurements were completed with a single row.

Scanning the beam was accomplished by using again the phasing property of the spiral, whereby a mechanical rotation of the spiral of 1° is equivalent to a change in phase of the far field of the spiral of 1°. By setting successive angular displacements in each element, the beam was scanned in the horizontal plane. The predicted angle of scan ($\theta$) is related to the phase difference between elements with half-wavelength spacing by the following relation:

$$\text{Phase difference (in degrees)} = 180 \sin \theta.$$

These differences were calculated and the spirals were then rotated the proper amount in order to obtain the scans shown in Figs. 4 through 9 (pp. 294–295).

## Results

The experimental program verified the theoretical considerations in general, and also pointed out some interesting characteristics of the system. For the broadside (0° scan) position of the main beam, the field on axis was found to be linear (approximately 40 db); the linearity is taken as the difference, in dbs, between the principal polarization of the main beam and the cross-polarized component. Fig. 4 shows this broadside position of the main beam for the principal polarization and for cross-polarization. The plane of polarization was found to be rotated 12° clockwise from the predicted horizontal position. The sidelobes were higher than the −25-db feedboard design which may be attributed to the fact that the feedboard was designed for a 25-db amplitude distribution looking into matched loads, while the introduction of a polarization grating between the spirals and the ground plane can conceivably alter this amplitude distribution. No attempt was made to further improve the sidelobe level, since the principal concern was the linearity of the polarization. As the beam was scanned out to ±45°, it was found that the linearity decreased with increasing scan angle (Figs. 5–9). Concomitant with this, the plane of the linear polarization rotated with increasing scan angle. This information is presented with each antenna pattern.

Another noticeable characteristic of this scanning array is the apparent increase in gain as the beam is scanned out to ±45°. This is due partially to the fact that the spiral placed $\lambda/8$ above the grid and grid placed $3\lambda/8$ above the ground plane had a dip, rather than a maximum, in the center of the pattern. Also, no Aquadag was placed on the outer turns of the spirals (as had been done in the eight-doublet array) to improve the axial ratio of the individual spirals. (Aquadag serves to attenuate any currents which might reach the outside edges of the spiral; if these currents were reflected from the ends of the spiral, they would radiate in a sense opposite to that of the currents traveling from the input terminals and increase the axial ratio. A high axial ratio serves to modulate both the amplitude and the phase of the radiation, the change in phase no longer being linear with spiral rotation and the amplitude no longer remaining invariant as the spiral rotates.) It might also be pointed out that the patterns of the individual spirals predicted better linearity for the large scan angles (not less than 14 db) than was actually obtained when the spirals were combined in the array.

## Conclusions

The principal advantage of the spiral-grating system is that it eliminates the need for a spiral-doublet configuration to obtain linear polarization and thus, in a two-dimensional array, allows for equal spacing of elements in both principal dimensions. Such a system could easily be scanned over a ±45° range without the introduction of extraneous main lobes (and even further, if the spiral pattern permitted). On the basis of the experimental work, this can be achieved if one is willing to accept linearity of approximately 10 db at the larger scan angles. The sidelobes obtained with the spiral-grating system were higher than those obtained with the eight-doublet array, but no further attempt was made to reduce the sidelobe level. Some other considerations are that the linearity drops off rather rapidly with increasing angle of scan and that the plane of polarization rotates as the beam is scanned. These factors may or may not be disadvantages, depending upon the particular application of the system.

The spiral-doublet array, in contrast with the spiral-grating system, produces a field that is linearly polarized and whose plane of polarization does not rotate with the scan angle. If doublets are used as the elements in an array, they can conceivably be spaced approximately $\lambda/\pi$ (at one frequency) in one dimension, so that a ±45° scan can easily be obtained. However, the minimum spacing in the other dimension will be $2\lambda/\pi(\sim 0.7\lambda)$, so that the beam can be scanned in this plane only ±30° before an extraneous main lobe is introduced. It seems, therefore, that the choice of either a spiral-grating system or a spiral doublet will be dictated by the particular use to which the system is to be put.

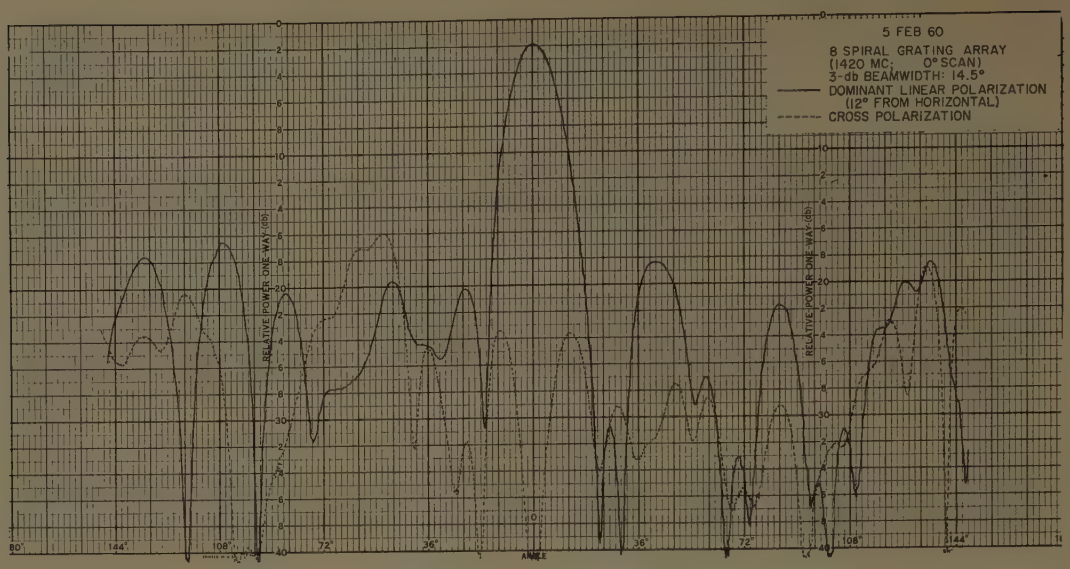



Fig. 4—Broadside pattern of spiral-grating array.

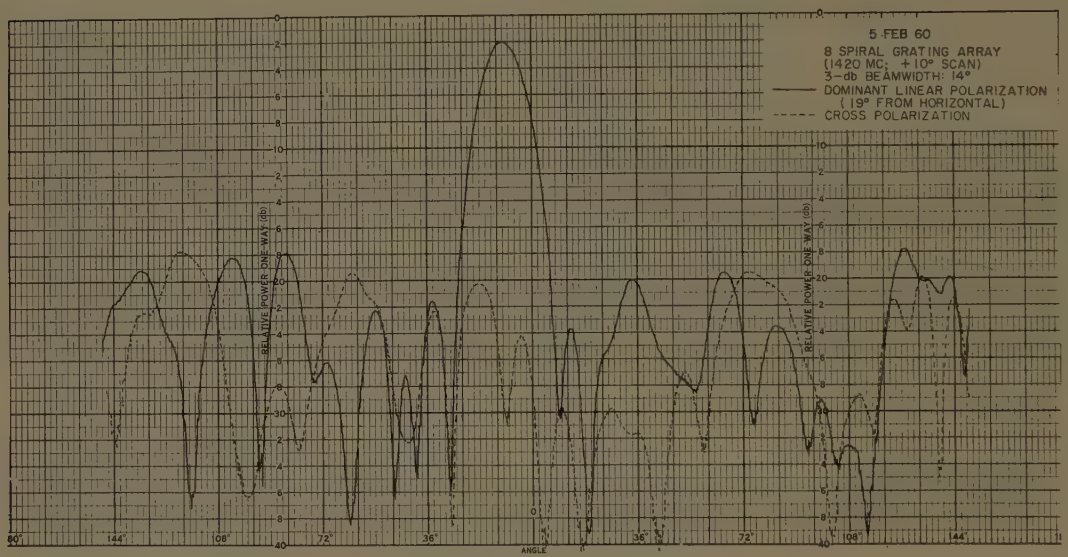



Fig. 5—Pattern of the spiral-grating array scanned +10°.

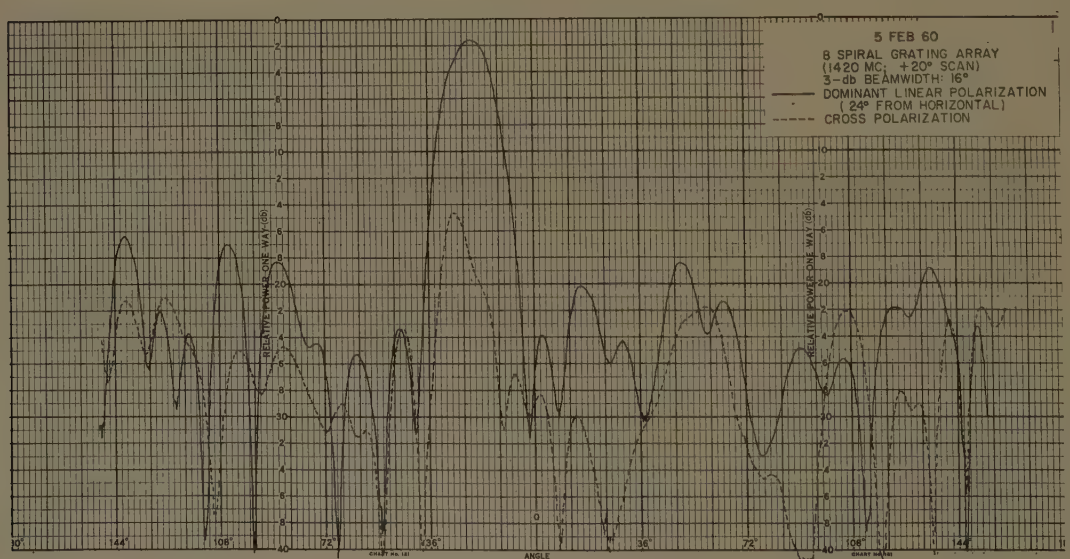



Fig. 6—Pattern of the spiral-grating array scanned +20°.

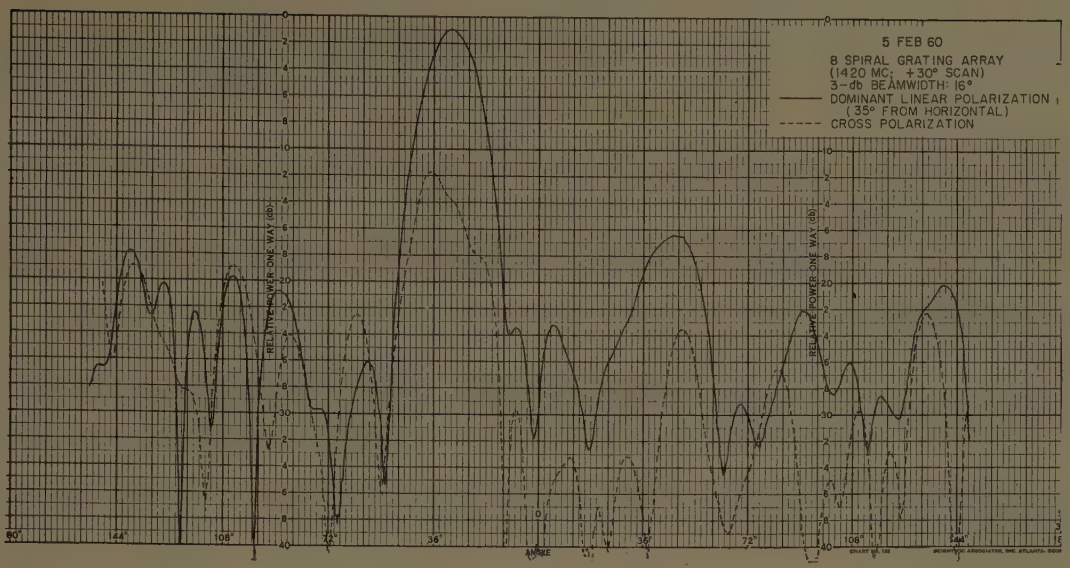



Fig. 7—Pattern of the spiral-grating array scanned +30°.

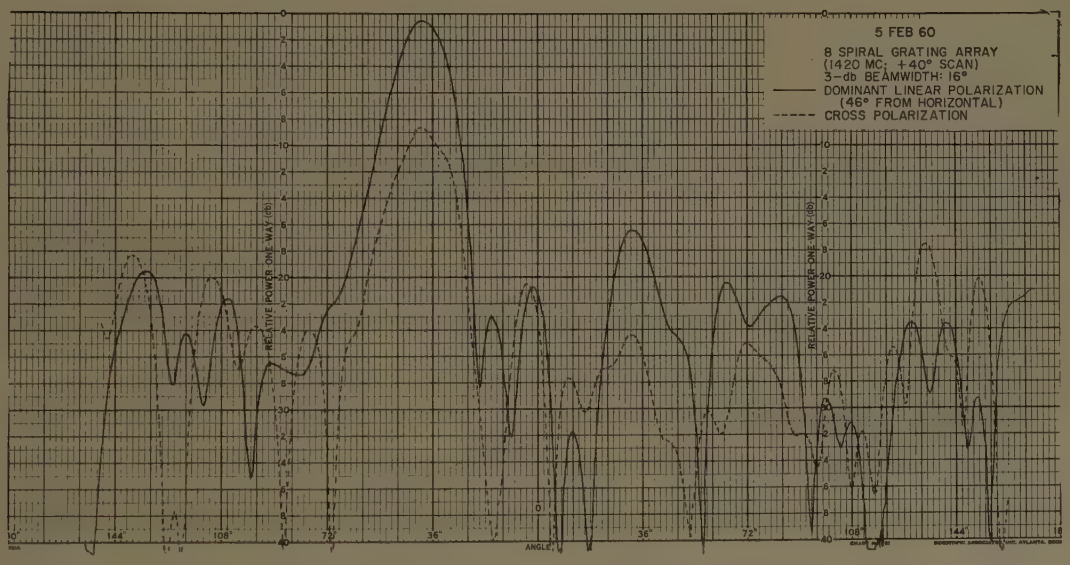



Fig. 8—Pattern of the spiral-grating array scanned +40°.

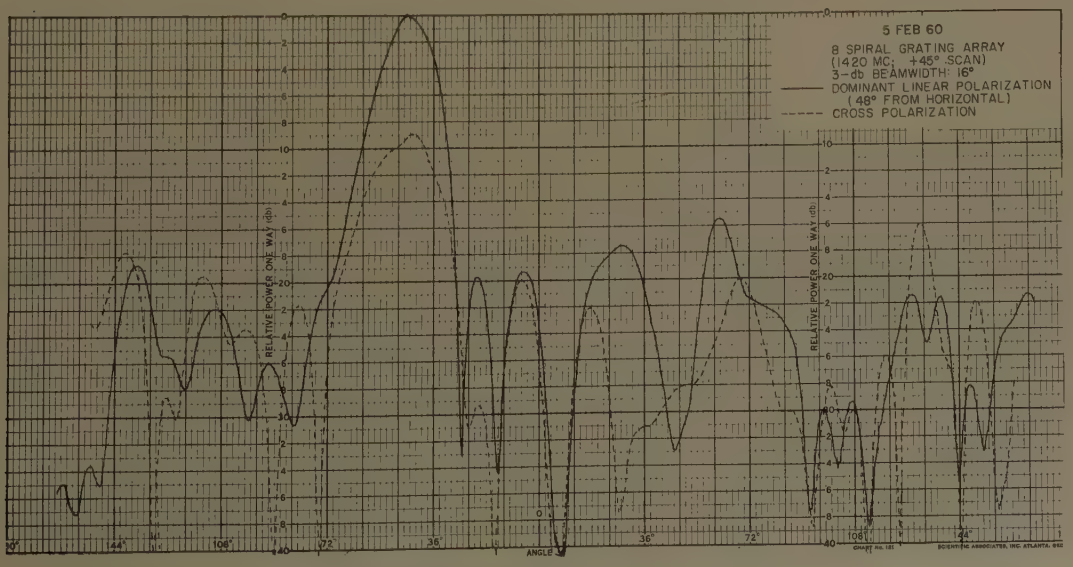



Fig. 9—Pattern of the spiral-grating array scanned +45°.

# Coupled Surface Waves and Broadside Arrays of End-Fire Antennas*

J. L. YEN†, MEMBER, IRE

*Summary*—**The concept of coupled surface waves is introduced to describe the action of element interaction in a broadside array of traveling-wave end-fire antennas. These coupled waves are formulated in terms of an interaction matrix, whose elements either can be determined experimentally, or, for some antenna structures, can be computed. Two methods of pattern evaluation based on two well-known approaches to end-fire antennas are then presented. The roles of interaction in the two methods turn out to be quite different, and the method which assumes radiation at the end of the antenna appears to be more favorable. A criterion on the over-all interaction effect on array pattern is introduced which indicates under what condition coupling can be neglected. Finally, the nature of patterns obtainable with different element spacings is examined.**

## Introduction

BROADSIDE arrays of traveling-wave end-fire antennas in one form or another such as helices, dielectric rods and Yagis have been in use for many years. In some instances, the antennas are separated by fairly large distances, while in others the spacings are small. Despite the variety of existing systems one factor common to most is that their proper functioning depends on negligible interaction between adjacent elements. The degree of interaction is usually determined experimentally from perturbations on pattern and impedance produced by coupling, and some qualitative measure is adopted to indicate its smallness. Only in one notable exception[1] is coupling actually utilized to achieve the desired pattern through proper adjustments. In view of the many advantages of broadside arrays of end-fire antennas, a better understanding of the role of element interaction is desirable. The purpose of this investigation is to explain the effect of interaction on array pattern from the coupled surface waves carried by the antenna system and to examine the types of patterns obtainable with such arrays. The results should be of assistance in exploiting element interaction in antenna systems and to predict when interaction can be neglected.

There are two well-known approximate approaches to the radiation of a traveling-wave end-fire antenna; one assumes the radiation of a surface wave along the antenna, while the other assumes the radiation of a surface wave at the radiating end of the antenna. To describe a broadside array of end-fire elements from either viewpoint, the concept of coupled surface waves on the antenna structures is introduced. These coupled waves are formulated in terms of an interaction matrix, whose elements either can be determined experimentally or, for some antenna structures, can be computed. Two methods of pattern evaluation based on the two approaches to end-fire antennas are then presented. The roles of interaction are different in the two methods and the one which assumes radiation at the end of the antenna appears more favorable. Finally the nature of patterns obtainable with different element spacing is examined.

## Radiation of a Broadside Array of End-Fire Antennas

Before we consider a broadside array of traveling-wave end-fire antennas, it is necessary to have a clear understanding of the individual elements which make up the array. The radiation of a traveling-wave end-fire antenna involves the diffraction by the finite antenna structure of an incident wave produced by the feed. Unfortunately, even the approximate solution of such a problem for a very simple geometric configuration involves considerable mathematical difficulties. A rigorous analysis of the general problem is out of the question. This leads to the use of crude approximate methods which can be more readily applied. The simplest and the most commonly employed approximation assumes the radiation of a surface wave along the antenna structure. Sometimes a direct feed radiation is added for higher accuracy. A general discussion of end-fire antennas of various structural types based on this method was given by Reynolds.[2] A second method capable of greater accuracy, but slightly more complicated, assumes the surface wave to travel without radiation along the antenna until it reaches the far end where it is diffracted, thereby producing radiation. The diffraction can be approximated by the radiation of an aperture field at the radiating end of the antenna whose distribution is simply that of the surface wave as shown by Brown and Spector.[3] The interference of this radiation and the direct feed radiation yields the radiation field of the end-fire antenna. Although the two methods assume different radiation mechanisms, computation of near fields of ideal discrete sources with traveling-wave excitation by Hyde[4] indicates that in certain situations the first method can also be interpreted as an aperture field

* Received by the PGAP, June 15, 1960; revised manuscript received, September 13, 1960. This investigation was partly supported by the Natl. Res. Council of Canada under Grant No. G 587.

† Dept. of Elec. Engrg., University of Toronto, Ont.

[1] H. W. Ehrenspeck and W. J. Kearns, "Two-dimensional end-fire array with increased gain and side lobe reduction," 1957 IRE WESCON CONVENTION RECORD, pt. 1, pp. 217–230.

[2] D. K. Reynolds, "Broadband traveling-wave antennas," 1957 IRE NATIONAL CONVENTION RECORD, pt. 1, pp. 99–107.

[3] J. Brown and J. D. Spector, "The radiating properties of end-fire aerials," *J. IEE*, vol. 104, pt. B, pp. 27–34; January, 1957.

[4] G. Hyde, "A Study of the Near Fields of Ideal End-Fire Antennas," M.A.Sc. thesis, Dept. of Elec. Engrg., University of Toronto, Ont.; October, 1959.

radiation interfering with the direct radiation from the feed. The computed near field distributions greatly resemble the experimental plots obtained by Ehrenspeck and Kearns.[1,5] Zucker and Kay[6] pointed out that in spite of neglecting turbulence in the surface field near the feed, the first method yields results that are useful in engineering design. In particular, it can be readily generalized to tapered or modulated antennas.[7] Thus, it appears that each method possesses certain advantages over the other.

Since both approaches provide useful information on traveling-wave end-fire antennas, it is worthwhile to examine broadside arrays of such elements from both viewpoints. Fortunately, the surface wave on the antenna structure is the common basis of the two methods. We can, therefore, begin with an investigation of surface waves on an antenna system. If the antenna structures are well separated, there can be no interaction between the surface waves on the different structures. The array pattern is then given by the product of the element pattern and the space factor determined by the complex feed amplitudes and the array configuration, regardless of which approach we use. When they are brought close together in parallel, the individual surface waves combine and redistribute into a set of coupled surface wave modes, each with a distinct amplitude distribution among the structures and each with a distinct phase velocity. If the spacing between antennas is not very close, the individual structures can be fed independently. The combined action of the feed system, however, is to excite in general, not one, but all the coupled modes. The coupled modes propagating along the system with different phase velocities result in a continuous redistribution of fields about the individual structures in the form of spatial beating which can be considered as modulation of the waves about the individual structures. The manner in which two guided waves interact has been discussed at length by Miller.[8] To calculate the radiation pattern using the first approach, we can either compute the radiation of the modulated waves on each structure first and then combine them to obtain the array pattern or, conversely, evaluate the radiation due to each coupled mode on the antenna system and then combine them. When the second approach is to be used, the distribution of the surface waves about the individual structures at the radiating end computed from the coupled modes can be used. Thus, knowing the coupled surface wave modes on the antenna system, the array pattern can be derived from the feed amplitudes of the individual elements using either method. It turns out that the two methods yield different array patterns. On the basis of the preliminary experimental data available, it appears that the second approach is more accurate.

For long antenna elements, because of the length available for interaction effect to accumulate, even very weak coupling at fairly wide spacing can perturb the array pattern considerably. However, knowing the coupled surface wave modes we can determine the feed amplitudes required for a given array pattern using the methods outlined above. At such spacings, if we further assume the surface waves are not reflected at the radiating end of the antenna system, the impedance of the feed on the individual elements will not be affected by the interaction. Thus, a feed system can be readily designed to give the desired feed amplitudes.

If the spacing between antennas becomes sufficiently close, we find that the individual elements can no longer be excited independently, *i.e.* each element is also excited by the feeds attached to the others. When this happens, the excitation of the coupled surface waves by the feed system becomes complicated. Moreover, this may lead to modifications in the feed impedances. To avoid these difficulties, such close spacings will be excluded from the following discussions.

Before presenting the two methods of pattern evaluation and their comparison, we shall begin with a discussion of the coupled surface waves on an antenna system. Without loss in generality, all subsequent discussions will be confined to one dimensional arrays with uniform spacing for simplicity in analysis.

## Coupled Surface Wave Modes

Consider $N$ identical structures, each of infinite length and each capable of supporting a surface wave without radiation in the absence of the others. In bringing them close together in parallel, the $N$ individual modes split into $N$ coupled modes. If the spacing between structures is large compared with the transverse dimensions of the structures, we expect the field distribution of a coupled mode in a transverse plane to be approximately given by the linear combination of fields centered about the individual structures, each with a distribution identical to that of the uncoupled surface wave. Let vector[9] $\mathfrak{r}$ denote position in a transverse plane and let the reference positions of the $N$ congruent structures be located at $\mathfrak{r}_n$, with spacing $D$ as shown in Fig. 1. The transverse electric field of the uncoupled mode on the $n$th structure propagating in the $+z$ direction with the phase constant

[5] H. W. Ehrenspeck and H. Poehler, "A new method for obtaining maximum gain from Yagi antennas," IRE Trans. on Antennas and Propagation, vol. AP-7, pp. 379–386; October, 1959.

[6] F. J. Zucker and A. L. Kay, "A Surface Wave Antenna Paradox," presented at the URSI Fall Meeting, Pennsylvania State University, University Park; October 22–29, 1958. Abstract appeared in IRE Trans. on Antennas and Propagation, vol. AP-7, p. 111; January, 1959.

[7] A. S. Thomas and F. J. Zucker, "Radiation from modulated surface wave structures—I," 1957 IRE National Convention Record, pt. 1, pp. 153–160.

[8] S. E. Miller, "Coupled wave theory and waveguide applications," *Bell Sys. Tech. J.*, vol. 33, pp. 661–719; May, 1954.

[9] We have adopted the convention that all German scripts are vectors, all lower case bold face Italics and German characters are column matrices with scalar and vector elements respectively. All capital bold face characters represent $N \times N$ matrices ***while v′ denotes the transpose of matrix.***

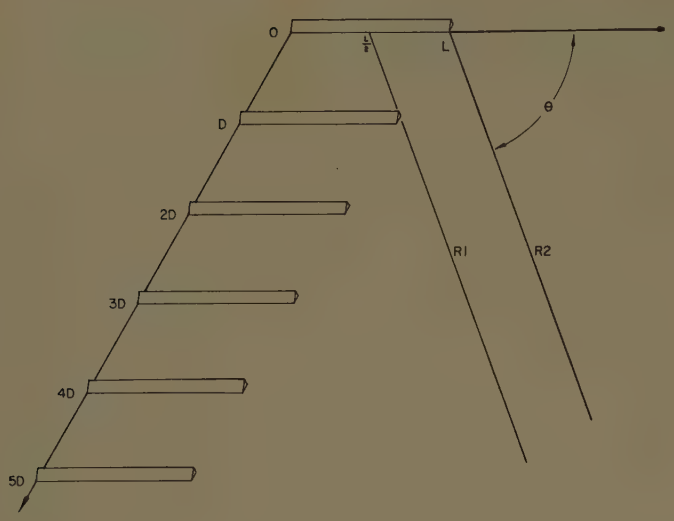



Fig. 1—A one-dimensional uniformly spaced array.

$\beta$ can be expressed as

$$\mathfrak{f}(\mathfrak{r} - \mathfrak{r}_n)e^{-j\beta z}. \tag{1}$$

Here, as well as in subsequent discussions, the time dependence $e^{j\omega t}$ is omitted. Let $(1+\Delta)\beta$ be the phase constant of a coupled mode. The transverse electric field of the mode propagating in the $+z$ direction can therefore be written in the following matrix form:

$$e(\mathfrak{r}, z) = \mathfrak{f}'ve^{-j(1+\Delta)\beta z}. \tag{2}$$

Here, $\mathfrak{f}'$ is a row matrix whose elements are the vector transverse electric fields in the transverse plane of the uncoupled modes on the $N$ structures as given by

$$\mathfrak{f}' = [\mathfrak{f}(\mathfrak{r} - \mathfrak{r}_1), \mathfrak{f}(\mathfrak{r} - \mathfrak{r}_2), \cdots, \mathfrak{f}(\mathfrak{r} - \mathfrak{r}_N)], \tag{3}$$

and where $v$ is the following column matrix whose elements are the complex amplitudes on the $N$ structures:

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_N \end{bmatrix}. \tag{4}$$

By going all the way to the eigen value problem which governs the system, a set of secular equations can be derived from which the perturbation in phase constant, $\Delta$, and the amplitudes matrix on the structures, $v$, for a coupled mode can be determined. Unfortunately, the equations are likely to be integral differential equations of a complicated form which cannot be solved using present techniques. However, under the above assumption of weak interaction, we expect the secular equations to be of the form

$$(K - \Delta 1)v = 0, \tag{5}$$

where $1$ is the unit diagonal matrix and $K$ is an $N \times N$ interaction matrix that describes the coupling effect. The off diagonal terms of $K$ should depend on the field at one structure due to the wave on another, while the diagonal terms should vanish. Although we cannot derive the secular equations, *e.g.*, (5) and the interaction matrix $K$ for all structural types even at reasonably large spacing, by using a multiple scattering approach,[10] we have been able to show, for an arbitrary structural type of small transverse dimensions supporting a dipole-type mode, that (5) is valid and that $K$ is of the form,

$$K = \begin{bmatrix} 0 & K(D) & \cdot & K(\overline{N-1}\,D) \\ K(D) & 0 & \cdot & K(\overline{N-2}\,D) \\ \cdot & \cdot & \cdot & \cdot \\ K(\overline{N-1}\,D) & K(\overline{N-2}\,D) & \cdot & 0 \end{bmatrix}, \tag{6}$$

where $K(D)$, the interaction coefficient, depends only on the spacing $D$ in terms of wavelength and the transverse propagation constant $q$ of the uncoupled mode defined as

$$q = \sqrt{\frac{\beta^2}{k^2} - 1} \tag{7}$$

according to

$$K(D) = \frac{1}{2}q^2\sqrt{\frac{2\pi}{qkD}}\,e^{-qkD}. \tag{8}$$

The transverse field distribution at a large distance $D$ from a structure of arbitrary shape supporting a surface wave of phase constant $\beta$ must behave as the asymptotic form of a Hankel function with imaginary argument, *i.e.* as

$$\sqrt{\frac{2}{\pi qkD}}\,e^{-qkD}. \tag{9}$$

Therefore, as expected, the interaction coefficient (8) is proportional to the field at one structure produced by the wave on another. For two special structural types, namely the dielectric rod[11] and the coated cylinder,[12] more accurate values of the interaction coefficient are available. In spite of the fact that the coated cylinder supports a circularly symmetric mode, both results agree favourably with (8). Therefore, unless more accurate values are available, (8) can be safely used. If the structure has a large traverse extension, $K(D)$ can be determined from a simple experiment to be described later.

Returning to (5) we see that $\Delta$ is an eigen value and $v$ the corresponding eigen function of the characteristic equation of $K$. From (8), observe that $K$ is real, hence there are $N$ real eigen values with orthogonal eigen functions that can be labelled as $\Delta^{(r)}$ and $v^{(r)}$, respectively. Real $\Delta^{(r)}$ implies all the coupled modes are propa-

[10] J. L. Yen, "Multiple scattering, propagation in periodic structures and coupled waves." To be published.

[11] M. F. Bracey, A. L. Cullen, E. F. T. Gillespie, and J. A. Staniforth, "Surface wave research in Sheffield," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-7, pp. S219–S226; December, 1959.

[12] A. A. Meyerhoff, "Interaction between surface wave transmission lines," PROC. IRE, vol. 40, pp. 1061–1068; September, 1952.

gating. Furthermore, $\boldsymbol{K}$ has zero diagonal terms, its trace vanishes, hence

$$\sum_{r=1}^{N} \Delta^{(r)} = 0. \tag{10}$$

We shall restrict our attention to situations where all the coupled modes are slow waves so that $(1+\Delta^{(r)})\beta > k$. Otherwise, some of the modes become fast waves, and the nature of the antenna system will be greatly altered. Substituting $\Delta^{(r)}$ and $v^{(r)}$ into (2), the transverse electric fields of the $N$ coupled modes, $e^{(r)}(x, z)$, are obtained.

For later discussion it is convenient to express the transverse electric fields of the coupled modes and their relation with the transverse fields of uncoupled modes on the individual structures in matrix form. Let the transverse electric field of the $N$ coupled modes be combined in a row matrix,

$$\mathbf{e}' = [e^{(1)}(x, z), e^{(2)}(x, z), \cdots, e^{(N)}(x, z)]. \tag{11}$$

The $N$ eigen-functions $v^{(r)}$ can be combined to form an $N \times N$ matrix $\boldsymbol{V}$, according to

$$\boldsymbol{V} = [v^{(1)}, v^{(2)}, \cdots, v^{(N)}]. \tag{12}$$

Since $v^{(r)}$ are orthogonal, they can be normalized so that $\boldsymbol{V}$ is orthogonal, *i.e.*, its transpose $\boldsymbol{V}'$ is also its inverse

$$\boldsymbol{V}'\boldsymbol{V} = 1. \tag{13}$$

The propagation of the coupled modes can be expressed in a diagonal matrix $\boldsymbol{P}(z)$ defined by

$$P(z) = e^{-j\beta z}\begin{bmatrix} e^{-j\Delta(n)\beta z} & 0 & \cdot & 0 \\ 0 & e^{-j\Delta(n)\beta z} & \cdot & 0 \\ \vdots & & & \\ 0 & 0 & \cdot & e^{-j\Delta(n)\beta z} \end{bmatrix}. \tag{14}$$

Using (2) it is seen that the coupled mode matrix (11) can be expressed in terms of the individual mode matrix $\mathbf{f}'$ according to

$$\mathbf{e}' = \mathbf{f}\boldsymbol{V}\boldsymbol{P}(z). \tag{15}$$

The value of the interaction coefficient $K(D)$ as given in (8) decreases exponentially with increasing distance $D$ between structures. Fig. 2 shows the variation of $K(D)$ vs $kD$ for various values of $q$. Often it is only necessary to consider coupling between immediate neighbors, thus only two off-diagonal terms in $\boldsymbol{K}$ need be retained. The characteristic equation, (10), is then readily solved, and the resulting $\Delta^{(r)}$ and $\boldsymbol{V}$ for $N=2$, 3, 4, 5 are shown in Fig. 3. Note that the two coupled modes for $N=2$ correspond to the well-known symmetric and antisymmetric modes. Also, from these examples it can be seen that the larger the number of structures $N$ in the system, the larger can the perturbation in phase constant $\Delta^{(r)}$ become. As we shall see later,

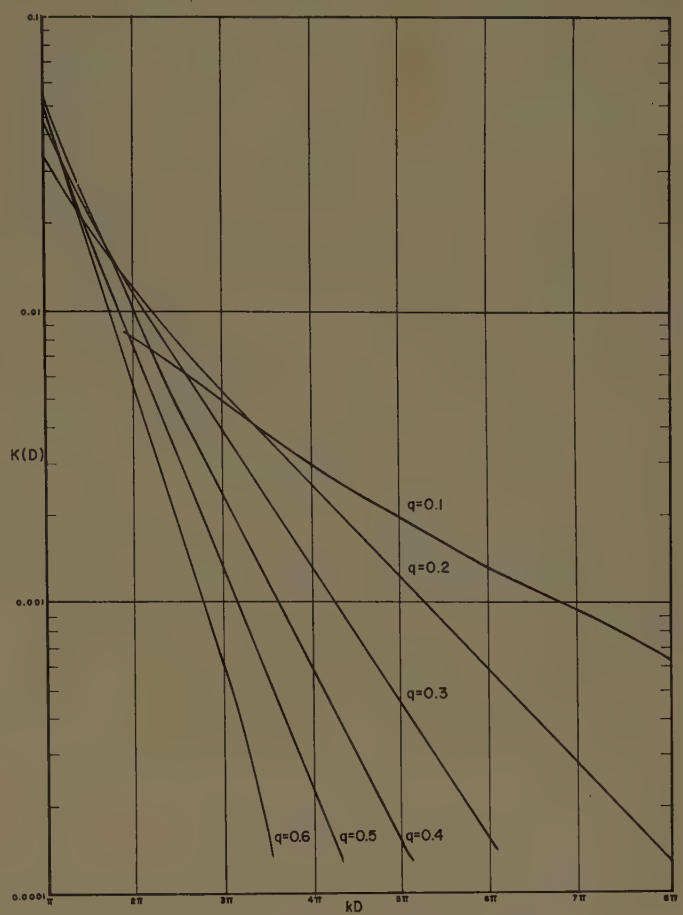



Fig. 2—Interaction coefficient $K(D)$.

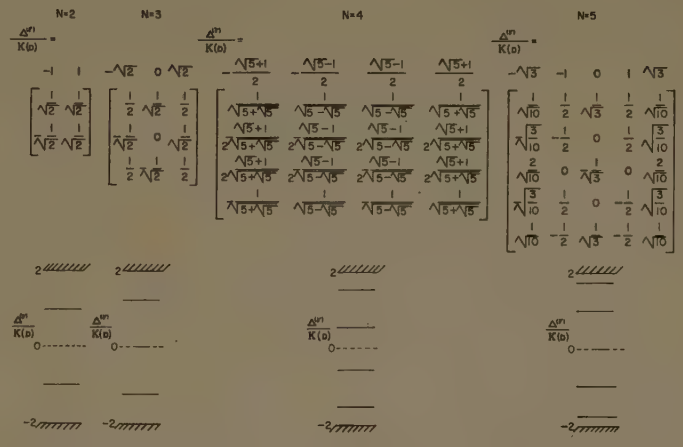



Fig. 3—$\Delta^{(r)}$ and $\boldsymbol{V}$ for interaction between adjacent elements only.

for an antenna system of finite length $L$, the effect of coupling on array pattern depends on $\Delta^{(r)}\beta L$. Thus, it would seem that the larger the number of structures, $N$, the more severe will be the overall coupling effect. Fortunately, $\Delta^{(r)}$ does not spread out over an infinite range as $N$ increases, only the density of the $N$ values of $\Delta^{(r)}$ increases in a finite band. This phenomenon is similar to the well-known band theory of conductors where the energy levels in a sample of finite number of atoms form discrete sets distributed within finite bands.

To see the existence of a finite band, consider a one-dimensional uniform array with $N$ infinite. The periodicity in the transverse direction requires the modes to satisfy Floquet's theorem, *i.e.*,

$$v_n^{(r)} = e^{-j\delta^{(r)}} v_{n-1}^{(r)}. \tag{16}$$

Assuming interaction between adjacent structures only, the characteristic equation, (5), reduces to the single equation.

$$\cos \delta^{(r)} = \frac{\Delta^{(r)}}{2K(D)}. \tag{17}$$

Since all the $v_n^{(r)}$'s are to remain finite, $\delta^{(r)}$ must be real. Thus we arrive at the finite allowed band

$$\left| \Delta^{(r)} \right| \leq 2K(D) \tag{18}$$

as illustrated in Fig. 3. A useful criterion in estimating the degree of coupling in an antenna system will be derived from this band.

In general, in an antenna system all the coupled modes are excited. Let the complex amplitudes of these coupled modes be expressed by the column matrix $\boldsymbol{a}$. The overall transverse electric field on the system is therefore

$$\mathfrak{E}(\mathrm{r}, z) = \mathfrak{e}'\boldsymbol{a} = \mathfrak{f}'\boldsymbol{V}\boldsymbol{P}(z)\boldsymbol{a}. \tag{19}$$

The differences in phase velocities will cause spatial beating between modes. This results in a continuous redistribution along the system of transverse field and energy distributions around the individual structures. It is only with very special excitations, that one coupled mode is excited. When this happens no redistribution of field and energy along the system will occur.

## Pattern Evaluation

Having determined the coupled surface wave modes, we now proceed to evaluate the array pattern using the two approximate methods of treating end-fire elements. For simplicity, only the pattern in the plane of the array will be discussed. We shall treat the two methods separately and then compare them. An experimental method for the determination of the interaction coefficient $K(D)$ will also be given.

The first approach assumes the coupled surface wave modes excited by the feed system to radiate along the antenna structure. The radiation of the $r$th coupled surface wave mode as given in (2), in the plane containing the array and of a particular polarization, is equivalent to that of the following traveling-wave current distribution on the system

$$v^{(r)} e^{-j(1+\Delta^{(r)})\beta z}. \tag{20}$$

The radiation of the current on each structure has its phase center located at the mid-point of the element ($z = L/2$). Using the coordinates $(R_1, \theta)$ as in Fig. 1, where $R_1$ is measured from the center of the first element, the radiation of a surface wave of unit amplitude along the first element is simply

$$C \int_0^L \frac{e^{-jkr}}{r} e^{-j(1+\Delta^{(r)})\beta z} dz = \frac{e^{-jkR_1}}{R_1} e^{-j(1+\Delta^{(r)})\beta(L/2)} F_1^{(r)}(\theta), \tag{21}$$

where a form factor $C$ is introduced which may vary with $\theta$ slowly in case of structures of large transverse dimensions. $F_1^{(r)}(\theta)$ is the well-known $\sin x/x$ pattern

$$F_1^{(r)}(\theta) = CL \frac{\sin x}{x}, \quad x = \left[(1 + \Delta^{(r)})\beta - k \cos\theta\right] \frac{L}{2}. \tag{22}$$

To obtain the total radiation of the $N$ travelling waves of (20) we need the space factor determined from the amplitudes of the elements and the array spacing $D$. Let $\boldsymbol{q}'$ be the following row matrix

$$\boldsymbol{q}' = \left[1, e^{jkD \sin\theta}, e^{j2kD \sin\theta}, \cdots, e^{j(N-1)kD \sin\theta}\right]. \tag{23}$$

It is readily seen that the radiation of the $r$th coupled mode is

$$\boldsymbol{q}'\boldsymbol{v}^{(r)} e^{-j(1+\Delta^{(r)})\beta(L/2)} F_1^{(r)}(\theta). \tag{24}$$

When all the coupled modes are excited with complex amplitude matrix $\boldsymbol{a}$ so that the total field on the antenna system is as given by (19), the radiation field $G_1(\theta)$ can be determined immediately from (24). If we introduce a diagonal coupled mode radiation matrix $\boldsymbol{F}_1(\theta)$ whose elements are $F_1^{(r)}(\theta)$ of (22), the result becomes

$$G_1(\theta) = \boldsymbol{q}'\boldsymbol{V}\boldsymbol{P}(L/2)\boldsymbol{F}_1(\theta)\boldsymbol{a}. \tag{25}$$

Assume that the elements can be excited independently with the complex feed amplitudes given by the column matrix $\boldsymbol{u}$. From (19), (15) and (14), there follows

$$\boldsymbol{u} = \boldsymbol{V}\boldsymbol{a}. \tag{26}$$

Using the orthogonal property of $\boldsymbol{V}$ as given in (13), (25) reduces to

$$G_1(\theta) = \boldsymbol{q}'\boldsymbol{V}\boldsymbol{P}(L/2)\boldsymbol{F}_1(\theta)\boldsymbol{V}'\boldsymbol{u}. \tag{27}$$

If there is considerable direct radiation from the feed system we have to take it into account. Let the first feed with unit amplitude contribute a direct radiation field $F_f(\theta)$ of the same polarization as in the above discussion but measured with respect to the feed end $z = 0$. The direct radiation of the entire feed system when measured with respect to the mid-point of the first element is then

$$G_{f1}(\theta) = e^{-jk(L/2)\cos\theta} F_f(\theta)\boldsymbol{q}'\boldsymbol{u}. \tag{28}$$

The over-all radiation pattern from the array is therefore

$$P_1(\theta) = |G_1(\theta) + G_{f1}(\theta)| = |\boldsymbol{q}'[\boldsymbol{VP}(L/2)\boldsymbol{F}_1(\theta)\boldsymbol{V}' + e^{-jk(1/2)\cos\theta}F_f(\theta)\mathbf{1}]\boldsymbol{u}|. \quad (29)$$

Note that if $\Delta^{(r)}\beta(L/2) \ll 1$, then $\boldsymbol{P}(L/2) \doteq e^{-j\beta(L/2)}\mathbf{1}$, $\boldsymbol{F}_1(\theta) \doteq F(\theta)\mathbf{1}$ where $F(\theta)$ is given by (22) with $\Delta^{(r)} = 0$. Thus, when $\Delta^{(r)}\beta L \ll 1$, (29) reduces to

$$P_1(\theta) \doteq |\boldsymbol{q}'[F(\theta)e^{-j\beta(L/2)}\mathbf{1} + e^{-jk(L/2)\cos\theta}F_f(\theta)\mathbf{1}]\boldsymbol{u}|. \quad (30)$$

The quantity in the bracket is the element pattern while $\boldsymbol{q}'\boldsymbol{u}$ is the space factor determined from the feed amplitudes and the array configuration. Since $\Delta^{(r)}$ lies within the finite band of (18), if coupling is to be neglected the following criterion must be satisfied:

$$2K(D)\beta L \ll 1. \quad (31)$$

The error in array pattern due to neglecting interaction is then approximately of the order of $2K(D)\beta L$.

Next let us consider the second approach which assumes the coupled surface wave modes excited by the feed system propagate to the radiating end of the antenna system and then radiate. The transverse electric field at the radiating end caused by these modes are then, from (19), (26) and (13),

$$\mathfrak{E}(\mathfrak{r}, L) = \boldsymbol{\mathfrak{e}}'\boldsymbol{a} = \boldsymbol{\mathfrak{f}}'\boldsymbol{VP}(L)\boldsymbol{a} = \boldsymbol{\mathfrak{f}}'\boldsymbol{VP}(L)\mathrm{V}'\boldsymbol{u}. \quad (32)$$

Let the radiation of a particular polarization caused by the aperture field $\mathfrak{e}(\mathfrak{r}-\mathfrak{r}_1)$ on the first antenna element measured with respect to the center of the aperture $(0, L)$ be

$$\frac{e^{-jkR_2}}{R_2}F_2(\theta). \quad (33)$$

The reason for using $R_2$ is that $\mathfrak{e}(\mathfrak{r}-\mathfrak{r}_1)$ is usually symmetrical and almost of uniform phase; hence the phase center of its radiation is at the center of the aperture. The total radiation field of the $N$-component aperture field distributions $\mathfrak{e}(\mathfrak{r}-\mathfrak{r}_n)$ with the complex amplitudes given by (32) is therefore,

$$G_2(\theta) = F_2(\theta)\boldsymbol{q}'\boldsymbol{VP}(L)\boldsymbol{V}'\boldsymbol{u}, \quad (34)$$

with $\boldsymbol{q}'$ as given in (23).

In addition to the radiation of the surface waves at the radiating end of the antenna system, the direct feed radiation as given by (28) has to be included. However, since it is necessary to measure the radiation field with respect to the radiating end of the first element, but not its mid-point, the contribution of the entire feed system in the same polarization is now

$$G_{f2}(\theta) = e^{-jkL\cos\theta}F_f(\theta)\boldsymbol{q}'\boldsymbol{u}. \quad (35)$$

The over-all radiation pattern from the array computed from the second method is finally

$$P_2(\theta) = |G_2(\theta) + G_{f2}(\theta)| = |\boldsymbol{q}'[F_2(\theta)\boldsymbol{VP}(L)\boldsymbol{V}' + e^{-jkL\cos\theta}F_f(\theta)\mathbf{1}]\boldsymbol{u}|. \quad (36)$$

Note that in order to neglect interaction, it is necessary that $\Delta^{(r)}\beta L \ll 1$, so that

$$P_2(\theta) \doteq |\boldsymbol{q}'[F_2(\theta)e^{-j\beta L}\mathbf{1} + F_f(\theta)e^{-jkL\cos\theta}\mathbf{1}]\boldsymbol{u}|, \quad (37)$$

where the quantity in the bracket is again the element pattern but determined from the second method. Introducing the allowed band of $\Delta^{(r)}$, the criterion of negligible coupling (31) is again applicable.

Comparing the over-all array pattern expressions $P_1(\theta)$ and $P_2(\theta)$ given in (28) and (36), the most pronounced difference lies in the appearance of matrix $\boldsymbol{P}(L/2)\boldsymbol{F}_1(\theta)$ in the first, and of matrix $\boldsymbol{P}(L)F_2(\theta)$ in the second. Although matrix $\boldsymbol{F}_1(\theta)$ depends on $\Delta^{(r)}\beta L$, its variation is of a different form and is less rapid compared with $e^{-j\Delta^{(r)}\beta(L/2)}$. Thus, the effect on the over-all pattern caused by element interaction predicted from the first method is less severe than that predicted on the basis of the second approach. To see this difference in more detail a simple example will be used, which will also provide an experimental determination of the interaction coefficient $K(D)$.

Consider a two-element array with the first element fed while the second unfed so that

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Note that we do not use equal excitations because only one coupled mode will then be excited. Using the coupled modes for $N=2$ as given in Fig. 3 the radiation pattern can be computed from (28) and (36). Assume that near the main beam we can neglect the direct feed radiation. The array pattern will have its maximum pushed to one side as in Fig. 4. The logarithmic derivatives of $P_1(\theta)$ and $P_2(\theta)$ at $\theta = 0$ are

$$\frac{d}{d\theta}\ln P_1(\theta)\Big]_{\theta=0} \doteq \left[1 + \frac{2K(D)\beta}{\beta - k} - kL\cot(\beta - k)\frac{L}{2}\right]\frac{kD}{2}\sin K(D)\beta L \quad (38)$$

$$\frac{d}{d\theta}\ln P_2(\theta)\Big]_{\theta=0} \doteq kD\sin K(D)\beta L. \quad (39)$$

The two approaches yield the factors $kD/z$ and $kD$, respectively, because the traveling-wave radiation along an element depends on the coupled surface waves at the mid-points of the element, while the radiation of an aperture field depends on the coupled surface wave at the radiating end of the element. The last two terms in the bracket of (38) account for the difference in patterns of the two coupled modes caused by different phase velocities. They are usually of the order of $K(D)\beta L$ so that the quantity in the bracket of (38) is approximately unity. The predicted value of (38) based on the first method is, therefore, less than that of (39) using the second approach. Note that from the measured pattern

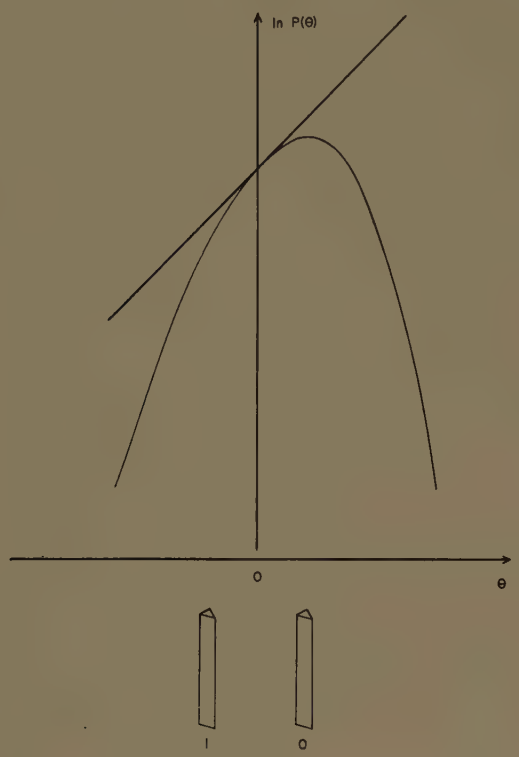



Fig. 4—Pushing of main beam caused by coupling in a two-element array.

of such a two-element array the interaction coefficient $K(D)$ can be determined, using either method.

Preliminary experiment[13] on long Yagi arrays seem to favour the second method. The values of interaction coefficient $K(D)$ obtained from measurement based on (39) are closer to the values computed from (8). Furthermore, the mechanism of radiation assumed by the second approach is more reasonable. Thus, it appears that in spite of the usefulness of the first approach in predicting the pattern of an end-fire antenna, it is less appropriate in describing the behaviour of a broadside array of such elements. However, this approach should be readily generalizable to broadside arrays of tapered or modulated elements because it describes the individual radiation of such elements quite adequately. A more definite assessment of the two methods, however, must await further experimental evidences.

## Effect of Spacing on Pattern

We shall now investigate the nature of patterns obtainable from a one dimensional uniformly spaced array in the plane containing the array. On the basis of the second method, the array pattern can be considered as the radiation of a virtual aperture field at the radiating end of the antenna system whose distribution includes both the coupled surface wave modes and the feed system direct radiation. The former dominates near the elements and decreases exponentially away from the elements, while the latter becomes important only far from the elements.

With very wide spacing, the array pattern is simply the product of the feed space factor $\boldsymbol{q}'\boldsymbol{u}$ and the element pattern as given by (37) or (30). It is well known that the space factor of a discrete source system with spacing $D>\lambda$ has more than one main beam in the form of a grating interferometer pattern as illustrated by the equal ripple pattern of Fig. 5. The angular intervals between the interferometer beams are small when the spacing is large. With sufficiently large $D$, a few of these beams may lie within the main beam of the element pattern, resulting in a multiple beam over-all pattern as in Fig. 5. Such an array is said to be of wide-spaced type whose virtual aperture field distribution exhibits wide gaps as shown in Fig. 5. Obviously, wide-spaced arrays are applicable only for special purposes.

When the spacing is decreased, the angular intervals between the interferometer beams in the feed space factor gradually increase. Soon, a spacing $D_0$ is reached at which the pair of first-order interferometer beams fall on the first minima of the element pattern at $\pm\theta_0$. The over-all pattern will therefore have all interferometer beams suppressed below the element side lobe leve. For an element pattern with sufficiently low side lobes, this results in a fan beam (pencil beam for a two-dimensional array) with unobjectionable side lobes using a minimum element density. An array of this type is, therefore, said to be optimum-spaced with optimum spacing $D_0$ given by

$$\lambda/D_0 = \sin\theta_0. \tag{40}$$

The main beam shape again depends on the illumination or the feed amplitudes in the usual manner. Fig. 6 illustrates the space factor, the element pattern and the over-all pattern of the array of Fig. 5, with spacing reduced to the optimum value given by (40), together with the virtual aperture field distribution which includes the direct feed radiation. Observe that fields of adjacent elements now overlap slightly resulting in an aperture field with gaps partially filled so that the interferometer beams are suppressed. In fact $D_0$ is simply the effective element aperture width defined by assuming a uniform distribution with the same first minima $\pm\theta_0$ as the actual element. With optimum spacing $D_0$ the aperture of the array is, therefore, completely illuminated over a total width $ND_0$.

Optimum-spaced arrays are economical fan or pencil beam antenna systems because a low element density is used. This not only reduces the total cost of the antenna elements but also simplifies the feed arrangement. Furthermore, if the effective element aperture width $D_0$ does not vary rapidly with frequency, the optimum-spaced condition can be satisfied for a reasonably wide

[13] D. L. Sengupta, D. D. Tang, and J. L. Yen, "Broadside Arrays of Long Yagis," presented at the Canadian IRE Convention, Toronto, Ont., October 7–9, 1959.

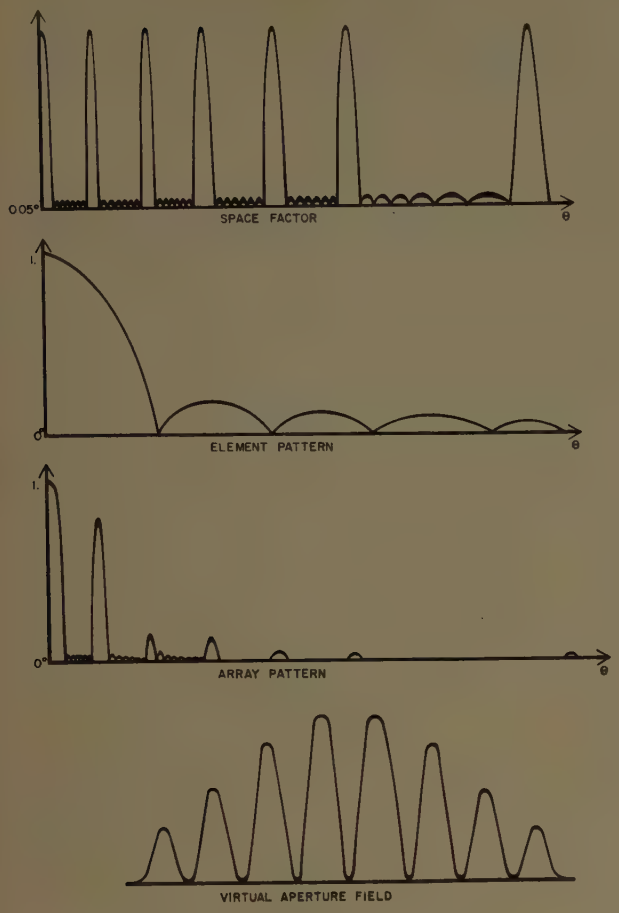



Fig. 5—A wide-spaced array.

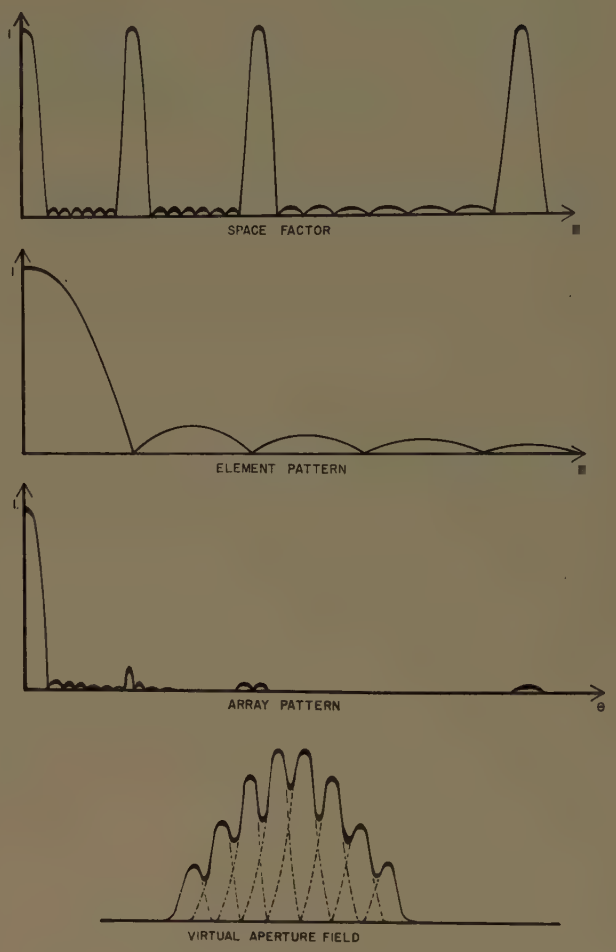



Fig. 6—An optimum-spaced array.

frequency band. Although the fields of adjacent elements overlap slightly as in Fig. 6, the field at one element because of its neighbours must still remain small because of the exponential behaviour. Therefore, condition (31) for negligible interaction is expected to remain valid for optimum-spaced arrays. This type of array has found wide applications.

When the spacing $D$ is reduced below optimum $D_0$, the fields and effective apertures of adjacent elements overlap so much that the system aperture width is only a fraction of $ND_0$. Such an array belongs to the close-spaced type with a beamwidth greater than that of an optimum-spaced array using the same number of elements unless a supergain technique is employed. In a nonsupergain close-spaced array the greater than optimum element density can be used to provide a finer control of the system aperture field so that no interferometer beam is evident as indicated in Fig. 7. A practical example is the dielectric rod array of Mickey and Chadwick[14] which consists of 17 $4\frac{1}{2}\lambda$ rods spaced at about $0.6\lambda$. In this array the coupling effect is present but seems to be neglected in the design of the array. In

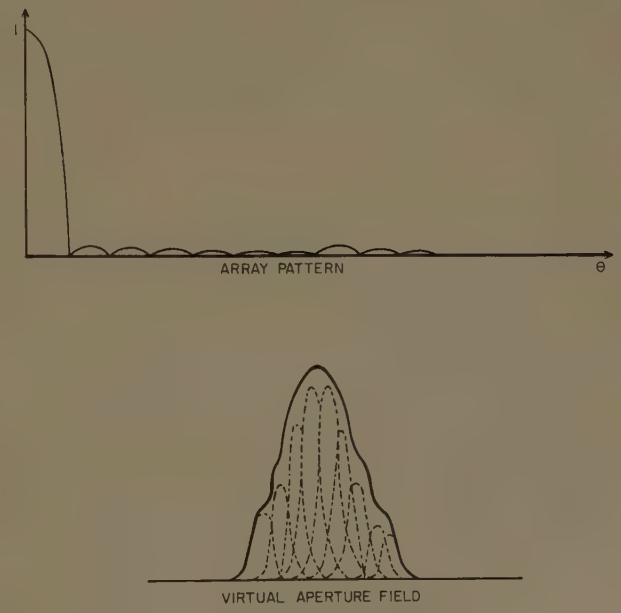



Fig. 7—A nonsupergain close-spaced array.

[14] L. W. Mickey and G. G. Chadwick, "Closely spaced high dielectric constant polyrod arrays," 1958 IRE National Convention Record, pt. 1, pp. 213–224.

contrast to this, the Yagi arrays of Ehrenspecks and Kearns' are also close-spaced, but the coupling is utilized for their operation. One of these Yagi arrays has a fed 6λ centre element and two 3.2λ parasites spaced at 0.66λ. The parasites are excited exclusively by coupled surface waves. The amount of coupling is controlled by the length and spacing of the parasites. When properly adjusted, a clean pattern of very low side lobes is obtained with beam width not appreciably narrower than that of the centre element. Antennas of this type are easily analyzed using the present theory. In fact, other antennas of this type should be derivable using the present analysis.

If it is desired to reduce the beamwidth of close-spaced array below that can be normally obtained with the total aperture width, it is necessary to adjust the feed amplitudes to produce a narrow-beam space factor as so aptly demonstrated by the French investigators[15] using nondirective elements. In an array of end-fire elements with narrow element beam the level of the space factor outside the element beam need not be kept very low. This results in a greater freedom in choosing the space factor, and hence, it should be easier to achieve supergain effect. It should be possible to design the feed system using the coupled surface waves. The spacing, of course, cannot be so close that the elements can no longer be independently excited.

[15] P. Aigrain, "Les antennes super-directives," *L'Onde Electrique*, vol. 33, pp. 51–54; February, 1952.

## Conclusion

We have presented two approximate methods of evaluating the pattern of a broadside array of end-fire antennas on the basis of the coupled surface wave modes on the antenna structures. The method assuming the radiation to occur at the radiating end of the array appears to be the more accurate of the two. However, the method assuming the radiation to occur along the antennas may be more readily generalized to tapered or modulated elements. Both approaches are approximate, applicable only for reasonably wide spacings so that the elements can be independently excited. Further experimental evidences are necessary to determine the accuracy of the predicted patterns. In spite of these deficiencies, the use of coupled surface waves should lead to a better understanding of the role of interaction between traveling-wave end-fire elements in a broadside array.

## Acknowledgment

The author wishes to thank Dr. George Sinclair for presenting a version of this paper at the URSI Spring Meeting, May, 1960, Washington, D. C.

# A Theory of Antenna Performance in Scatter-Type Reception*

S. STEIN†, SENIOR MEMBER, IRE, AND D. E. JOHANSEN†, MEMBER, IRE

*Summary*—**Many of the diversity techniques currently proposed or in use for scatter reception may be viewed as utilizing antenna systems with several adjustable parameters—typically two to four—which are continually varied in response to fluctuations in the incident waveform, so as to maximize received power. An investigation is described which defines an ideal upper bound on performance of a linear variable-parameter antenna system in a scatter-type reception situation. The important results are 1) a standard for evaluation of antenna performance, and 2) deeper physical insight into the problem of antenna operation in scatter-type reception.**

**A reception model is used in which the receiving antenna views a set of radio sources on a hypothetical celestial sphere. For each such set of sources possible, a "matched" antenna excitation exists which maximizes received power. Assuming random fluctuations in the radio source consistent with those observed in the transhorizon phenomenon, statistical properties of the received power are derived for the optimized antenna. Similar properties are then investigated for standard diversity systems. A significant "gap" is found to exist between the performance of existing or contemplated systems and the theoretical optimum.**

* Received by the PGAP, July 22, 1960; revised manuscript received, October 31, 1960. The research reported in this paper was performed at Hermes Electronics Co., Cambridge, Mass., under Contract AF 19(604)-5214 with the Electronics Research Directorate, AF Cambridge Res. Ctr., Air Res. and Dev. Command.

† Appl. Res. Lab., Sylvania Electronic Systems, a Div. of Sylvania Electric Products, Inc., Waltham, Mass.

## I. Introduction

IN recent years, a significant effort has been directed towards developing and improving the performance of transhorizon communications at frequencies of VHF and above, *i.e.*, so-called tropospheric scatter. In order to circumvent the deleterious effects of characteristically weak, fading received fields, increasing use has been made of such techniques as large radiated power, large antennas, and special redundancy techniques usually categorized as diversity

The present study focuses attention on possibilities for further improvements in antenna performance without additional increases in antenna size. The aim of this investigation has been to achieve an increased understanding of antenna operation under conditions typified by scatter propagation, and of the nature of possible optimal antenna systems for such applications.[1] The results presented in this paper provide a useful framework for evaluating the performance and relative efficiencies of antenna systems for tropospheric scatter, including, *e.g.*, space and angle diversity [2]–[5].

## II. Antenna Reception Equations[2]

We assume a linearly polarized antenna system characterized by a planar aperture distribution, an assumption consistent with most UHF and microwave large antenna design. From the well-known Fourier transform relation between aperture excitation and radiation pattern, and by applying the rigorous electromagnetic reciprocity theorem, we can obtain the following basic statements:

In transmitting, let the complex-valued distribution of electric field in the aperture plane be $E_0(x, y)$, assumed to vanish outside the antenna aperture $S$. Assume that $E_0(x, y)$ gives rise to a radiated field whose complex-valued angular pattern is $R(\alpha, \beta)$, where $\alpha$, $\beta$ are the direction cosines of a far-field point with respect to the coordinate axes $x$, $y$ of the aperture plane. Then if the antenna is a relatively narrow-beam (high-gain) antenna, and is nonsuperdirective, the following transform relationships hold (with conveniently chosen proportionality constants):

$$R(\alpha, \beta) = \frac{1}{\lambda}\int\int_{-\infty}^{\infty} E_0(x, y) \exp\left[-\frac{i2\pi}{\lambda}(x\alpha + y\beta)\right] dxdy$$

$$= \frac{1}{\lambda}\int_S E_0(x, y) \exp\left[-\frac{i2\pi}{\lambda}(x\alpha + y\beta)\right] dxdy \quad (1)$$

$$E_0(x, y) = \frac{1}{\lambda}\int\int_{-\infty}^{\infty} R(\alpha, \beta) \exp\left[\frac{i2\pi}{\lambda}(x\alpha + y\beta)\right] d\alpha d\beta,$$

where $\lambda$ = wavelength and $S$ is the domain covered by the physical antenna aperture. Further, the total radiated power is

$$P_0 = \int\int_{-\infty}^{\infty} |R(\alpha, \beta)|^2 d\alpha d\beta = \int\int_{-\infty}^{\infty} |E_0(x, y)|^2 dxdy$$

$$= \int_S |E_0(x, y)|^2 dxdy, \quad (2)$$

where the bars indicate magnitudes of complex numbers.

Next, suppose that we are receiving signals from a distribution of sources in the far field. This distribution will be described by a complex "voltage" distribution $f(\alpha, \beta)$ defined over a distant "celestial" sphere, with our antenna located at the center of this sphere, and being used for reception. (The scale on $f(\alpha, \beta)$ is always interpreted indirectly through antenna measurements and need be of no concern here.) The distribution of sources corresponds qualitatively to the various multipath rays which appear to "originate" within the so-called "common volume" of the transhorizon propagation phenomenon. In addition, the sources are all assumed to be coherent; *i.e.*, even though the RF amplitudes and phases of these signals are randomized, a basic waveform which is identical for all is associated with each source signal. Thus, $f(\alpha, \beta)$ may be thought of as describing RF carrier amplitude and phase for signals emanating from the point $(\alpha, \beta)$.[3] The power detected at the receiving antenna terminals, if we use the same antenna characterized earlier by its transmitting properties, is then readily shown to be

$$P_R = \frac{1}{P_0}\left|\int\int_{-\infty}^{\infty} R(\alpha, \beta) f(\alpha, \beta) d\alpha d\beta\right|^2$$

$$= \frac{1}{P_0}\left|\int\int_{-\infty}^{\infty} E_0(x, y) F_0(x, y) dxdy\right|^2 \quad (3)$$

$$= \frac{1}{P_0}\left|\int_S E_0(x, y) F_0(x, y) dxdy\right|^2$$

where we have additionally defined $F_0(x, y)$, the field on the aperture plane due to the distant sources,

$$F_0(x, y) = \frac{1}{\lambda}\int\int_{-\infty}^{\infty} f(\alpha, \beta) \exp\left[-\frac{i2\pi}{\lambda}(x\alpha + y\beta)\right] d\alpha d\beta. \quad (4)$$

Eq. (3) bears a direct analogy to the circuit theory problem of the signal-noise ratio at the output of a filter. This suggests the existence of an antenna analog to the "matched filter." It is, indeed, readily shown that for a fixed source distribution $f(\alpha, \beta)$, the maximum available received power is obtained by adjusting the antenna feed so that, over the aperture, the feed excitation factor $E_0(x, y)$ is a conjugate match to the incident field,

$$E_0(x, y) = KF^*(x, y), \quad (5)$$

where $K$ is an arbitrary constant; and that this maxi-

[1] The only related previous work appears to be that of J. Granlund [1]. Granlund, however, sought an essentially fixed parameter system for maximizing mean received power, whereas our concentration is on minimizing the major deleterious effects due to the (statistically rare) deep fades.

[2] To restrict this paper in length, it has been necessary to delete any amplified discussion of the basic antenna equations. Such a detailed discussion, including both the "non-superdirective" assumptions referred to in this section, and a complete discussion of the reciprocity relationships, is available in [6].

[3] We will here treat $f(\alpha, \beta)$ as a surface density (in a sense which will be apparent later). However, all results are fully general and can be applied to a collection of discrete sources by, for example, treating $f(\alpha, \beta)$ as a collection of appropriate delta functions with complex coefficients.

mum available received power is

$$(P_R)_{\max} = \int_S |F(x, y)|^2 dxdy, \tag{6}$$

where $S$ is the antenna aperture.[4]

It is relevant, as a side remark here, to note an important connection of (6) with the diversity concept. The derivation holds even if the aperture is comprised of a sum of disconnected subapertures, in which case the maximum received power can be written

$$P_R = \sum_i \int_{S_i} |F(x, y)|^2 dxdy = \sum_i P_{R_i}. \tag{7}$$

But this is just the output from a maximal ratio diversity combiner [7] operating on the outputs of the several optimized subapertures if they are associated with equal noise levels. Carrying this argument a little further, it becomes apparent that the optimum antenna defined by (6) can be regarded as a diversity system having essentially an *infinite* number of degrees of freedom.

The result in (6) gives the maximum power that can be received with a non-superdirective antenna and an arbitrary source distribution, and has been derived on the basis that the intensity of the source distribution is nonfluctuating. In the sections to follow, we will deal with source distributions which fluctuate in time (as does the apparent source in scatter reception), with the details of the fluctuations described statistically. However, we now argue that it is valid to assume as our ideal antenna system one with variable parameters that are appropriately controlled to follow the source fluctuations with a negligibly small time lag. This is physically reasonable since fading rates in UHF transhorizon propagation are in the order of 1–10 cps, while, for example, diversity combiners are in operation with time constants of the order of a couple of milliseconds.

Thus, we can use (6) as a description of the instantaneous received power for our conceptual optimum antenna, and the problem is to relate, through this equation, the statistics of $P_R$ to those assumed for the fluctuations of the sources. Similarly, we will utilize (3) to determine the instantaneous $P_R$ for appropriate "non-optimum" antennas, and subsequently again the related statistics.

## III. Performance of the Optimum Antenna

In this section, we apply the results of Section II, and evaluate the appropriate statistics for the fluctuating power output of the optimum receiving antenna, for a specific "extended" source configuration fluctuating statistically in time in all its elements.

[4] It is well known that only half the incident power flux can actually be abstracted from the field. But since our assumed source function is to be interpreted only in terms of antenna measurements, one should not become unduly disturbed about the "missing" numerical factor in (6).

### A. Choice of Source Distribution

Much of the early phenomenological interpretation of transhorizon propagation assumed that the signal arrives over essentially a continuum of directions of arrival, with randomly varying associated strengths. More recently, it has been argued [8] that the scattered signal may be comprised of a few discrete "rays" (such as might arrive by specular reflection from a few moving tilted layers in the atmosphere). Unfortunately, the available experimental data is not sufficient to resolve the question satisfactorily [9]. In view of this uncertainty, it seems desirable to evaluate the performance for both discrete and continuous source models. In this paper, we will limit the discussion to the more general continuous distribution case. Limited computations which have also been already carried out for discrete-source models have yielded the same qualitative conclusions as will be discussed here [6].

### B. Formulation as an Eigenvalue Problem

We first introduce convenient dimensionless variables to denote positions on the aperture plane, $\xi = x/\lambda$, $\eta = y/\lambda$. In terms of these new variables, the incident electric field is written $F(\xi, \eta) = F(x/\lambda, y/\lambda) = F_0(x, y)$.

The distribution of apparent sources on the celestial sphere may be described by an ensemble of complex source functions $f(\alpha, \beta)$ with a corresponding ensemble of images on the aperture plane given, from (4), by

$$F(\xi, \eta) = \frac{1}{\lambda} \int\int_{-\infty}^{\infty} f(\alpha, \beta) \exp\left[-i2\pi(\xi\alpha + \eta\beta)\right] d\alpha d\beta. \tag{8}$$

We assume that the fluctuations in signal carrier emitted from any point $(\alpha, \beta)$ are uncorrelated with the fluctuations in signal carrier emitted from every other point,[5] and we describe the spatial statistics of $f(\alpha, \beta)$ by the correlation functions of its real and imaginary components. Thus, we first write

$$f(\alpha, \beta) = \bar{f}(\alpha, \beta) + i\hat{f}(\alpha, \beta).$$

In addition, we assume symmetry in the statistics of the quadrature components,[6] with expectations given by

$$E[\bar{f}(\alpha_1, \beta_1)\bar{f}(\alpha_2, \beta_2)] = E[\hat{f}(\alpha_1, \beta_1)\hat{f}(\alpha_2, \beta_2)] = \frac{\lambda^2}{2} g(\alpha_1, \beta_1)\delta(\alpha_1 - \alpha_2)\delta(\beta_1 - \beta_2) \tag{9}$$

$$E[\bar{f}(\alpha_1, \beta_1)\hat{f}(\alpha_2, \beta_2)] = 0.$$

[5] The assumption of zero correlation from point-to-point is a valid mathematical approximation for our purposes to the more physically likely situation at the ranges and beamwidths involved in practice, that the "correlation distance" on the celestial sphere is very small compared to the distances required for significant changes in the antenna beam pattern.

[6] Some insight into the generality of this assumption may be obtained from [10].

From (8) and (9), it is readily shown that

$$E[|F(\xi,\eta)|2] = \int\int_{-\infty}^{\infty} g(\alpha,\beta)d\alpha d\beta. \quad (10)$$

Hence $g(\alpha, \beta)$ can be interpreted as the contribution-per-unit solid angle on the celestial sphere to the average power density in the incident wave front.

The integral form in (8) for $F(\xi, \eta)$ and the assumed narrow autocorrelation for $f(\alpha, \beta)$ imply via the central limit theorem that the real and imaginary values of $F(\xi, \eta)$, taken at arbitrary points on $(\xi, \eta)$, obey a multivariate Gaussian distribution. The statistics of $F(\xi, \eta)$ are thus completely specified by its complex autocorrelation function.

$$R_0(\xi_1,\xi_2;\eta_1,\eta_2) = E[F(\xi_1,\eta_1)F^*(\xi_2,\eta_2)] \quad (11)$$

$$= \int\int_{-\infty}^{\infty} d\alpha d\beta g(\alpha,\beta)\exp\{-i2\pi[\alpha(\xi_1-\xi_2)+\beta(\eta_1-\eta_2)]\}.$$

The right-hand side of (11) is a function only of the two differences $t_1=\xi_1-\xi_2$ and $t_2=\eta_1-\eta_2$. Hence it is convenient to define

$$R(\xi_1-\xi_2,\eta_1-\eta_2) = \lambda^2 R_0(\xi_1,\xi_2;\eta_1,\eta_2), \quad (12)$$

where we have also introduced a normalization with respect to the wavelength $\lambda$ such that

$$(1/\lambda^2)R(\xi_1-\xi_2;\eta_1-\eta_2)$$

is the actual autocorrelation of $F$. We thus obtain the form

$$R(t_1,t_2) = \lambda^2\int\int_{-\infty}^{\infty} d\alpha d\beta g(\alpha,\beta)\exp\left\{-\frac{i2\pi}{\lambda}(\alpha t_1+\beta t_2)\right\}. \quad (13)$$

It is noted that $R(t_1, t_2)=R^*(-t_1, -t_2)$, *i.e.*, $R$ is Hermitian.

We next turn our attention to the probability distribution for the functional given by (6),

$$P_R = \lambda^2\int_S |F(\xi,\eta)|^2 d\xi d\eta, \quad (14)$$

which describes the power received by a continuously optimized antenna having the aperture $S$ on the $(\xi, \eta)$ plane. The distribution of such a functional has been determined by Kac and Siegert [11], [12] for the one-dimensional case where $S$ is a line segment and $F(t)$ is a function of the single variable $t$. The extension to two dimensions is straightforward [6]. The result is that the Laplace transform of the probability density function for $P_R$ is given by

$$L(z) = E[\exp\{-zP_R\}] = \prod_r (1+z/\sigma_r)^{-1}, \quad (15)$$

where the $\sigma_r$ are the eigenvalues of the integral equation,

$$\phi_r(\xi,\eta) = \sigma_r\int_S R(\xi-\xi_1,\eta-\eta_1)\phi_r(\xi_1,\eta_1)d\xi_1 d\eta_1. \quad (16)$$

### C. Numerical Solution

The integral equation (16) has been solved analytically for a circular aperture and a class of rotationally symmetric sources [6], [13]. These rotationally symmetric source functions $g(\alpha, \beta)$ are given as the ratio of two polynomials, as follows:

$$g(\alpha,\beta) = \frac{1}{\lambda^2}\frac{N[-4\pi^2(\alpha^2+\beta^2)]}{M[-4\pi^2(\alpha^2+\beta^2)]}, \quad (17)$$

where the polynomials $N(x)$ and $M(x)$ are of degree $n<m$, respectively. Substitution into (13) immediately provides the required kernel $R(t_1, t_2)$. Assuming the aperture to be circular with a physical radius $r$, the eigenvalues $\sigma$ for (16) are obtained [6], [13] as the denumerable infinity of roots of the set of $m\times m$ determinants:

$$E^{(k)}(\sigma) = \det\{E_{lj}^{(k)}(\sigma)\};\ k=0,\pm1,\pm2,\cdots\pm\infty, \quad (18)$$

whose $(lj)$th elements are

$$E_{lj}^{(k)}(\sigma) = \frac{\sigma}{h_j^2-\gamma_l^2}\left[h_j^{-|k|+1}I_{k+1}\left(h_j\frac{r}{\lambda}\right)K_k\left(\gamma_l\frac{r}{\lambda}\right) + \gamma_l h_j^{-|k|}I_k\left(h_j\frac{r}{\lambda}\right)K_{k+1}\left(\gamma_l\frac{r}{\lambda}\right)\right], \quad (19)$$

where $I_k(z)$ and $K_k(z)$ are the usual modified Bessel functions. The constants $\gamma_l^2$ are the roots of the polynomial $M(x)$, and the $h_j^2$ are the roots of the polynomial

$$H_\sigma(x) = M(x) - \sigma N(x). \quad (20)$$

Treating the $h_j^2$ as functions of $\sigma$ [via (20)], the determinantal equations $E^{(k)}(\sigma)=0$ are readily solved for the eigenvalues $\sigma_{kn}$, either graphically or by using standard iterative numerical procedures.

We have computed the eigenvalues $\sigma_{kn}$ for the simplest physically meaningful source distribution of the class considered,

$$N(x) = \frac{Q\lambda^2}{\pi r_s^2}; \qquad M(x) = \left(\frac{x}{8\pi r_s^2}\right)^2 + 1. \quad (21)$$

The various coefficients in (21) have been chosen so that first [by (10), (17), and (21)],

$$E[|F(\xi,\eta)|2] = \frac{Q}{\pi r_s^2}\int\int_{-\infty}^{\infty}\frac{d\alpha d\beta}{\left[\frac{-4\pi^2(\alpha^2+\beta^2)}{8\pi r_s^2}\right]^2} = Q \quad (22)$$

is interpretable as the average power density of the electromagnetic wave incident on the aperture plane, and second, the intensity of the source at its center is [by (17)]

$$g(0,0) = \frac{Q}{\pi r_s^2}. \quad (23)$$

It is clear from (22) and (23) that $r_s$ may be interpreted as the "radius" of the equivalent *uniform* source which also has a total flux density $Q$ on the aperture plane.[7]

In the later physical interpretations of our results, it will be convenient to introduce the parameter $r_s/r_a$ where $r_a$ is the radius of a typical beam "normally" radiated by the given aperture. Specifically, $r_a$ is taken to correspond to the radius of the equivalent uniform beam having the same maximum radiated power density, and the same total radiated power as does a uniformly excited aperture of radius $r$. For the wavelength $\lambda$, it is found that

$$r_a = \lambda/\pi r. \tag{24}$$

The eigenvalues computed for the source (21) are shown in Fig. 1 (opposite), plotted against $r_s/r_a$.[8] The eigenvalue $\sigma_{kn}$ is defined as the $n$th zero of the $k$th determinantal equation $E^{(k)}(\sigma)$. Actually, many more than fifteen eigenvalues are defined by Fig. 1, since for $k \geq 1$ the roots $_{\sigma kn}$ must be counted twice.

### D. Error Rate Computations

In the region of low signal levels of primary interest in determining the performance of most communications systems, the probability density function for $P_R$ can be accurately approximated via a saddle point integration performed on the integral defining the inverse transform of $L(z)$. However, a relation of more direct physical interest is available. Namely, the probability of error for matched filter detection of FSK signals in white noise is related to the instantaneous signal power $P_R$ by [14]

$$e(P_R) = \frac{1}{2}\exp\left[-\frac{P_R}{2\gamma N}\right], \tag{25}$$

where $\gamma$ is the information rate (bits per second) and $N$ the noise power density in the data channel, referred to the same point in the system as $P_R$. For a fluctuating $P_R$, with probability density function $w(P_R)$, the average error rate is

$$E = E[e(P_R)] = \frac{1}{2}\int_0^\infty \exp\left[-\frac{P_R}{2\gamma N}\right] w(P_R)\,dP_R. \tag{26}$$

Since in VHF and UHF communications the significant noise source generally lies within the receiver, the noise power density in (26) can be treated as a constant. Comparing with (15), (26) is recognized immediately as

$$E = \frac{1}{2}\prod_{kn}\left[1 + \frac{q}{\sigma_{kn}\dfrac{Q\lambda^2}{r_s^2}}\right]^{-1}, \tag{27}$$

where

$$q = \frac{Q\lambda^2}{2\gamma N r_s^2} \tag{28}$$

is a normalized average signal-to-noise ratio, and the $\sigma_{kn}(Q\lambda^2/r_s^2)$ are just the "dimensionless" eigenvalues plotted in Fig. 1.

Using (27) and the eigenvalues already determined, the error rates have been computed for a range of values of $r_s/r_a$. These error rate curves are plotted as the solid lines ($M=1$) in Figs. 2–4 for $r_s/r_a = 0.19$, 0.63, and 3.14, respectively.

## IV. Performance of Some Typical Nonoptimum Antennas

We next compare the potentialities represented by the performance of an optimized antenna with the performance of typical diversity arrays. We consider both a space diversity array and an angle diversity antenna receiving from the same source as does the optimum system. We will also generalize the optimum antenna results to a spaced array of optimum antennas separated sufficiently from one another so that the fading of the several signals is statistically independent. The results of all three computations are also plotted in Figs. 2–4.

### A. Space Diversity Array

A common diversity arrangement consists of a spaced array of two or more antennas, separated so that fading of the signal on each may be considered statistically independent of that on any other.

We assume that each diversity antenna has a Gaussian beam shape[9] with equivalent uniform radius $r_a$. The average power received on any particular beam is obtained by taking the average value of (3) with the aid of (9) and with the Gaussian beam shape substituted for $R(\alpha, \beta)$. Numerical integration of the resulting integrals yields $E(P_R) = 0.15\ Q\lambda^2/r_s^2$, $0.019\ Q\lambda^2/r_s^2$, and $0.0052\ Q\lambda^2/r_s^2$ for the average power received on a nonoptimized circular aperture when the ratio of source-to-beam width is $r_s/r_a = 0.19$, 0.63, and 3.14, respectively.

For the case of Rayleigh fading, knowledge of the received power is sufficient to define the fading distribution. The Laplace transform for the fading distribution of the received signal is

$$L_1(z) = [1 + zE(P_R)] - 1. \tag{29}$$

[7] The mean square width of the source (21) does not exist. Hence, the radius of the equivalent uniform source is used here as the appropriate measure of source width.

[8] This choice of parameter coincides with physical intuition which demands that except for a trivial renormalization on power, merely increasing the radius $r$ of the antenna aperture (and therefore decreasing the achievable beamwidth) should be completely compensated by proportionally decreasing the source radius $r_s$.

[9] The Gaussian antenna pattern is an approximation to the "Bessel function" pattern which results when a circular aperture is illuminated by a "flat" distribution of electric field $E(\xi, \eta) = 1$. Use of the Gaussian shape is justified by its closeness to the Bessel function for signals arriving not too far from beam axis, and by the fact that lack of sidelobes more accurately approximates the design ideal.

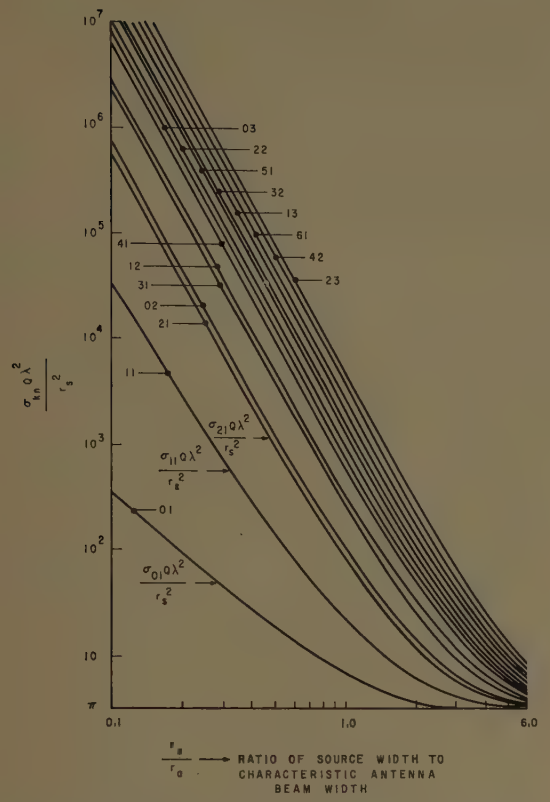



Fig. 1—Eigenvalues for optimum antenna receiving from extended source.

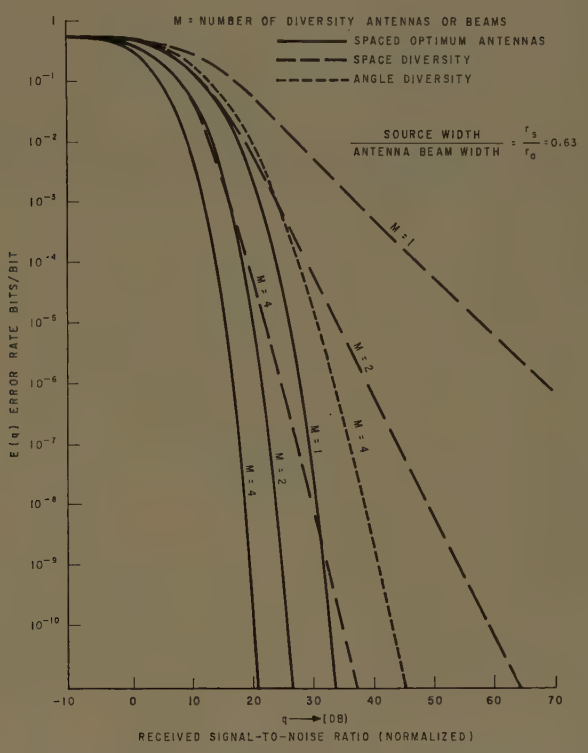



Fig. 3—Comparison of optimum antenna with diversity arrays.

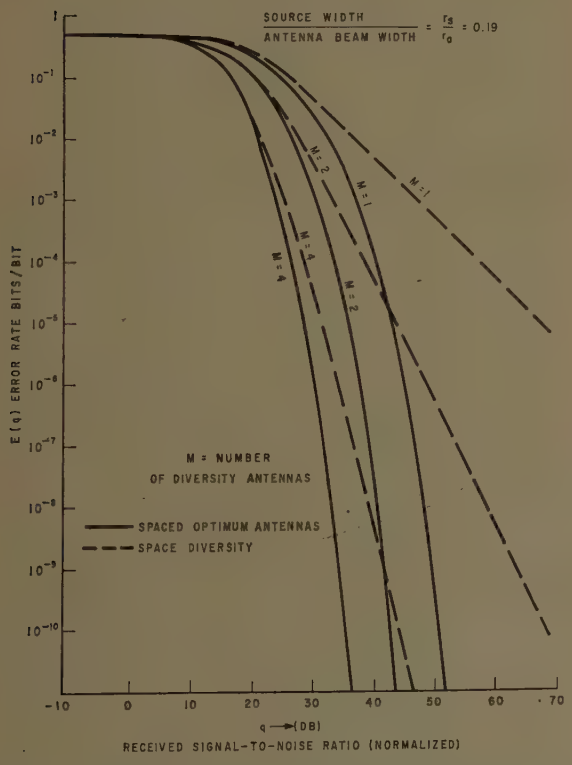



Fig. 2—Comparison of optimum antenna with diversity arrays.

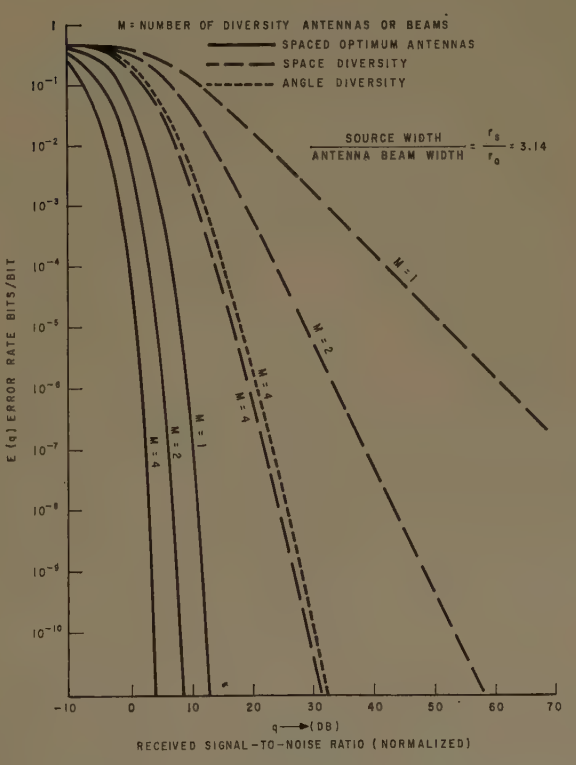



Fig. 4—Comparison of optimum antenna with diversity arrays.

Assuming independence among the several diversity signals, the Laplace transform of the combined output (maximal ratio diversity) is just the product

$$L_M(z) = L_1(z)M = \frac{1}{[1 + zE(P_R)]M}, \tag{30}$$

where $M$ is the order of diversity. By analogy to the development of (27) the corresponding error rate is then obtained as

$$E_M(q) = \frac{1}{2} \frac{1}{\left[1 + \frac{qE(P_R)}{Q\lambda^2/r_s^2}\right]^M}, \tag{31}$$

where the parameter $q$ is the same normalized average received carrier-to-noise ratio defined previously in (28). Curves for $M = 1, 2, 4$, are given in Figs. 2–4.

### B. Angle Diversity

Error rate curves for systems using four-fold angle diversity[10] and having source to beam-width ratios $r_s/r_a = 0.63$ and 3.14, are plotted in Figs. 3 and 4, respectively. The derivation of these curves is presented elsewhere [6]. In view of the possible lack of generality of the model[10] used there, with respect to real systems of angle diversity, these lengthy derivations are omitted here. It suffices to add that in general it is not physically possible to place the usual kind of feed horns any closer than that which will provide beam overlap at about the 3-db points. This physical restriction—that the angle diversity beam patterns overlap outside their 3-db points—implies that the ratio of source width to effective antenna beamwidth (*i.e.*, $r_s/r_a$) should not be taken very small, since otherwise the angular spread between diversity beams will be so large that sufficient coverage of the source by all diversity beams will not be possible. For this reason curves for the case $r_s/r_a = 0.19$ have not been included.

### C. Spaced Array of Optimum Antennas

Assuming that the separate structures of the array are sufficiently separated, fading on the individual optimum antennas within a spaced array will be statistically independent. Maximal ratio combining of the outputs (equivalent to the entire array being optimized) then produces an error rate which is derived by analogy to the development of (30), as follows:

$$E_M^{(0)}(q) = \tfrac{1}{2}[2E_1^{(0)}(q)]M, \tag{32}$$

where the functions $E_1^{(0)}(q)$ are the error rates computed earlier for a single optimum aperture, and $M$ is again the number of circular structures in the array. These results have also been plotted in Figs. 2–4.

## V. Conclusions

Figs. 2–4 present a comparison between the predicted performance of optimum antennas and the performance of existing or proposed systems of space and angle diversity. The proper measure of performance for these systems is obtained by comparing power requirements for a specified error rate performance. Several conclusions seem warranted:

1) A single optimum antenna offers a performance comparable to that of a standard space-diversity array of similar sized apertures. For the smaller of the three apertures considered, $r_s/r_a = 0.19$, the average received power requirement for an optimum single aperture is seen to be the same as that for two-fold space diversity, at an error rate of $10^{-5}$ bits per bit. A similar statement can be made for the $r_s/r_a = 0.63$ case. However, in the case of the largest of the three antennas, $r_s/r_a = 3.14$, the apparent superiority of the optimum antenna over even fourfold space diversity should be regarded with care, since a large part of the apparent superiority is due to the so-called "aperture-to-medium coupling loss," resulting when a fixed parameter aperture of large radius is receiving from a relatively large source. Thus the comparison for $r_s/r_a = 3.14$ appears to show not so much the ineffectiveness of diversity vs the optimum, but rather a demonstration that the *optimum effectively avoids aperture-medium coupling loss.*
2) Further significant improvements should be possible in the aperture utilization of existing antenna systems: Thus over the useful range of system error rates ($10^{-4}$ to $10^{-7}$ bits/bit) the optimum antenna outperforms a fourfold angle diversity system of the same aperture size by roughly 5–15 db.[11] Though only a fourfold angle diversity system has been analyzed, it is doubted that consideration of higher-order angle diversity systems would significantly change these conclusions, because of difficulties inherent in placing feed horns any closer together than that which will cause an overlap at the 3-db points. As a further example, it is seen that for a ratio of source width to effective antenna beamwidth of $r_s/r_a = 0.63$, optimization of a fourfold space diversity array will provide a power advantage of 6–7 db at the $10^{-5}$ error rate level, over the nonoptimized system.

We may also note that the performance gap between the "standard" diversity arrays and the optimum becomes even greater if we discuss error rates in the range $10^{-7}$

[10] The particular type of angle diversity discussed is as follows: The scatter volume is illuminated by a single transmitting antenna so as to provide an effective source as seen by the receiving structure, approximated by (17) and (21). The receiving structure consists of a single dish, with multiple feeds providing four independent overlapping receiving patterns with a symmetrical projection on the celestial sphere. The outputs from each of the four receiving patterns are combined in a maximal ratio combiner. The angular spread of the resulting pattern is, in every case, selected to provide best performance, *i.e.*, minimum power requirements for a specified error rate.

[11] These figures are based only on the two larger of the three aperture sizes considered.

to $10^{-10}$ bits per bit; requirements approaching the latter seem to be evident in applications of the near future, if not in those already existing.

A final note may again be made of the fact that all computations have been performed for the idealized situation involving a circular aperture and a circularly symmetric source distribution having a "radially" symmetric intensity distribution of the form $g(x) = (cx^4+1)^{-1}$. Although the absolute performance figures predicted in this context obviously will not remain valid for less symmetric situations, the relative differences existing among system types and the qualitative features of individual systems are expected to be preserved.

One may conclude that further research is warranted for filling the "gap" shown by our comparisons. An indication of why the optimum antenna performs better than more standard systems is offered by Figs. 2–4. The error rate curves for the diversity systems asymptotically approach a straight line, whereas those for the optimum bend continually downwards towards the vertical. It may be shown that the slopes of the straight line asymptotes are proportional to the order of diversity or number of degrees of freedom, while the continual bending of the optimum is related to the earlier remark (see the conclusion of Section II) that it is equivalent to an infinite number of degrees of freedom. Clearly, the direction for filling the gap lies in antennas which also equivalently have a large number of degrees of freedom. It is hoped to report further research in that direction in the near future.

## References


[1] J. Granlund, "Topics in the Design of Antennas for Scatter," Lincoln Lab., MIT, Lexington, Mass., Tech. Rept. 135, November 23, 1956.

[2] R. Bolgiano, Jr., N. H. Bryant, and W. E. Gordon, "Diversity Reception in Scatter Communications with Emphasis on Angle Diversity," School of Elec. Engrg., Cornell University, Ithaca, N. Y., Res. Rept. EE 359; January, 1958.

[3] J. H. Vogelman, J. L. Ryerson, and M. H. Bickelhaupt, "Tropospheric scatter system using angle diversity," PROC. IRE, vol. 47, pp. 688–696; May, 1959.

[4] A. B. Crawford, D. C. Hogg, and W. H. Kummer, "Studies in tropospheric propagation beyond the horizon," *Bell Sys. Tech. J.*, vol. 38, pp. 1067–1178; September, 1959.

[5] C. Chu, A. W. Wren, and J. LaRue, "Evaluation of the Pincushion System," University of Michigan, Ann Arbor, Rept. 2872-I-T, RADC-TN-50-60; February, 1960.

[6] S. Stein, D. Johansen, and A. Starr, "Final Report on Theory of Antenna Performance in Scatter Type Reception," Hermes Electronics Co., Cambridge, Mass., AF Cambridge Res. Center, Tech. Rept. 59-191; September 30, 1959.

[7] D. G. Brennan, "Linear diversity combining techniques," PROC. IRE, vol. 47, pp. 1075–1102; June, 1959.

[8] A. T. Waterman, Jr., "The Mechanism of Transhorizon Propagation: Layers vs Turbulence," presented at the URSI-IRE Joint Meeting, Washington, D. C.; May, 1959.

[9] D. E. Johansen, "On the Interpretation of Beam Sweep Profiles," presented at the URSI-IRE Fall Meeting, San Diego, Calif., October 19–21, 1959. See also [6].

[10] W. B. Davenport and W. L. Root, "Random Signals and Noise," McGraw-Hill Book Co., Inc., New York, N. Y.; 1958.

[11] M. Kac and A. J. F. Siegert, "An explicit representation of a stationary Gaussian process," *Annals Math. Stat.*, vol. 18, pp. 438–442; April, 1947.

[12] M. Kac and A. J. F. Siegert, "On the theory of noise in radio receivers with square-law detectors," *J. Appl. Phys.*, vol. 18, pp. 383–397; April, 1947.

[13] D. Johansen, "A Two-Dimensional Kac-Siegert Problem and Associated Integral Equation," submitted to IRE TRANS. ON INFORMATION THEORY. See also [6].

[14] S. Reiger, "Error probabilities of binary data transmission systems in the presence of random noise," 1953 IRE CONVENTION RECORD, pt. 8, pp. 72–79.


# Correction

Klaus Walther, author of "Reflection Factor of Gradual-Transition Absorbers for Electromagnetic and Acoustic Waves," which appeared on pages 608–621 of the November, 1960, issue of these TRANSACTIONS, has called the following to the attention of the *Editor*.

1) Eq. (46) on page 614 should read:

$$\bar{y}_\square = \bar{\sigma} l; \text{ where: } (\bar{\sigma})^2 = \sigma \int_0^1 \xi\sigma(\xi) \int_\xi^1 \sigma(\xi')d\xi' d\xi.$$

2) In the sentence following (47) on page 614, the first inequality should read: $\bar{y}_\square < y_c$; the second inequality should be: $\bar{y}_\square \geq y_c$.

3) In the caption of Fig. 18 on page 619, the expression $g(x) = s(x) - jt(x) \equiv 1$ should be corrected to: $g(x) = s(x) - jt(x)$; $f(x) \equiv 1$.

# communications

## A Scattering Measurement Technique*

King and Scharfman have described a technique for measuring the radar cross section of objects by modulating the position of the object.[1] The inherent simplicity and accuracy of this technique even within the confines of a small laboratory very much favor its use for many scattering problems, although it has been restricted to the measurement of cross sections of rotationally symmetrical obstacles.

A relatively simple mechanical improvisation on the experimental apparatus previously used with this technique has allowed us to measure successfully the scattering characteristics of three-dimensional, arbitrarily shaped objects. In response to King's suggestion, this improvisation is reported here in the hope that it may significantly extend the value of what is already an excellent experimental arrangement.

Generally the method consists of measuring the amplitude of the audio signal that corresponds in frequency to the Doppler shift caused by a relative motion between an antenna and target. A component of proper motion necessary for the measurement is realized by revolving the scatterer in a large circle. Since this motion, however, results in a continuous change of target aspect, the technique has been considered to be valid only for targets exhibiting rotational symmetry. The design pictured in Fig. 1 overcomes this limitation. Gear 1 is fixed with respect to the observer. As the horizontal arm is rotated about the drive shaft axis, the chain drives gear 2 so that the styrofoam column B rotates about its own vertical axis with respect to the horizontal arm as it simultaneously revolves about the drive-shaft axis. The rotation of the target effected by the chain drive exactly counters the rotation that would otherwise occur with respect to an observer during the course of target revolution. The aspect of the target thus remains fixed at all times with respect to the observer.[2] Since

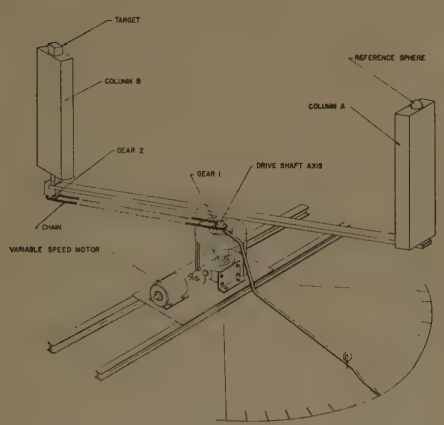



Fig. 1—Apparatus used for making scattering measurements of rotationally unsymmetrical targets.

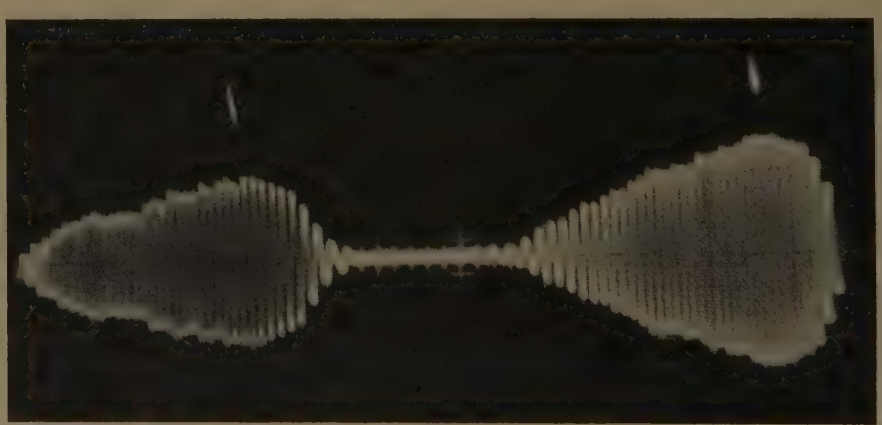
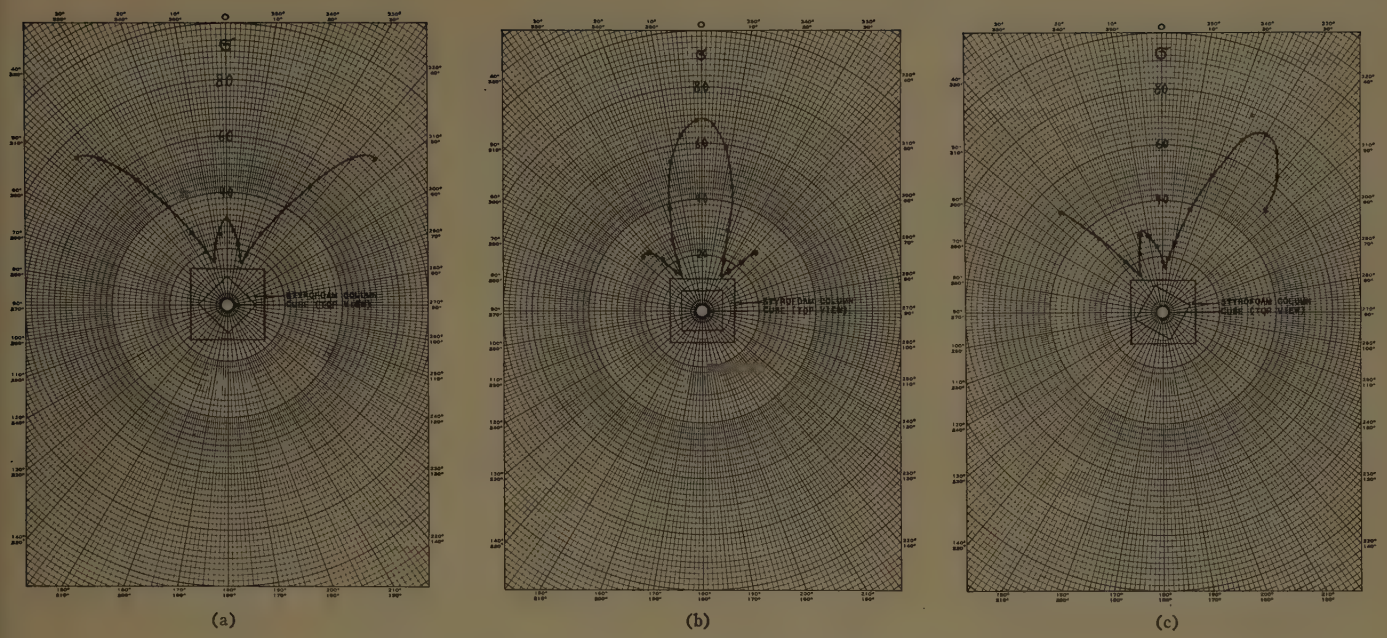


Fig. 2—Oscilloscope display of the return signals from a reference sphere and target. The pulse marks the point at which the amplitude of the signal should be read. The first signal is from a 2½-inch steel sphere and the second from a 2½-inch lucite cube.

* Received by the PGAP, December 12, 1960.

[1] For details of the measurement theory, see D. D. King and H. Scharfman, "Antenna-scattering measurements by modulation of scatterer," PROC. IRE, vol. 42, pp. 854–858; May, 1954.

[2] This neglects the change in aspect with respect to an observer due to the lateral displacement associated with the circular path of the target. The extent to which this may be neglected depends on the distance between observer and target.



(a) (b) (c)

Fig. 3—Backscattering cross section ($\sigma$) vs angular position of the cube. Aspect of the cube and styrofoam support with respect to the measured pattern is indicated.

column A supports a rotationally symmetrical standard scatterer, it may be fastened directly to the horizontal arm. Oscilloscope display of the return signal together with a positioning pulse (see Fig. 2) from a microswitch located on the rotating member will provide that both backscattering and off-angle scattering can be measured. The primary advantage of an oscilloscope display is the ease with which spurious noise components entering or being generated within the system may be detected. For example, modulation of the amplifier power supply by the ac motor, which was found to be a serious source of error, may have otherwise been unnoticed. The major disadvantage of the oscilloscope is of course the difficulty in reading accurately and the time required to obtain a number of readings in order to minimize the error. Although it has not been tried, a peak reading digital voltmeter commanded by the positioning pulse might overcome these objections.

In order to determine the linearity and general reliability of the system, the backscattered signals from a set of six steel spheres ranging in size from one to four inches in diameter were measured. A comparison of the experimental results with a theoretical curve (the only reference readily available at present) has indicated that the experimental error is at least as small as the error in reading the curve. All measurements were reproducible within $\pm 5$ per cent, and by averaging for many readings the error was reduced to $\pm$one per cent. This indicates that the primary source of the error is in reading the oscilloscope. Generally it may be said that the inherent accuracy of the modulation technique is not impaired by the new arrangement.

In order to measure the backscattering cross sections of rotationally unsymmetrical targets as a function of their angular position, a moveable boom was fastened to gear 1 (see Fig. 1). Manual movement of this boom results in a repositioning of the chain and, consequently, in a corresponding angular change in column and target. An angular scale associated with the boom facilitates the making of an accurate and reproducible change in target aspect without stopping its rotation. In this manner the angular position of a 2½-inch lucite cube was changed in 5° increments up to 45° to either side of a plane of symmetry and the return signals recorded. It was found that the physically symmetrical properties of the cube would provide a measure of error resulting from inaccurate positioning of the target with respect to the incident electric vector and due to the influence of the styrofoam supports. This was discovered when the first set of data taken with the cube was found to be unsymmetrical and when changes in the means of target support and compensation for the centrifugally-induced target misalignment resulted in the recording of almost perfectly symmetrical patterns. (See Fig. 3).

Since the support columns are not symmetrical, any influence they have may be expected to show up in the form of an unsymmetrical measurement pattern, as was first obtained. It was found that a reduction in size of the styrofoam support in the vicinity of the target improved the symmetry of the pattern. It may therefore be concluded that the support columns should be as small as possible; just how small can be determined by the above method. In other words, if the pattern is perfectly symmetrical and the support columns are unsymmetrical, then the columns can have had no influence. Fig. 4 is a composite of Fig. 3(a)–(c), and represents the errors resulting from positioning of the target [(a)–(c) of Fig. 3 each involved a new position of the cube], the effect of the styrofoam column, and the

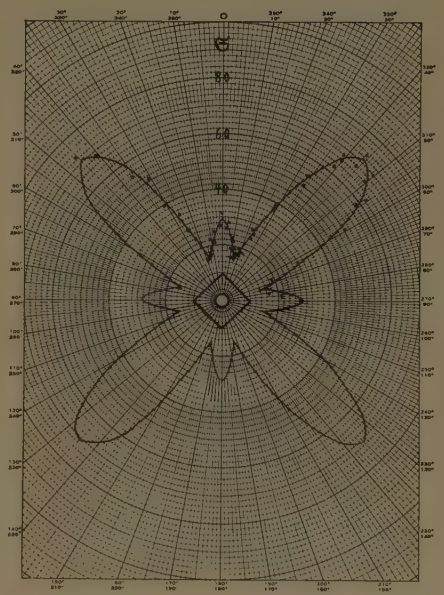


Fig. 4—Superposition of the points shown in Fig. 3(a)–(c), indicating the deviation of the measured points from an average symmetrical curve.

systematic error in reading the oscilloscope. It should be pointed out that the value of $\sigma$ is proportional to the square of the voltage ratios read on the oscilloscope.

These measurements were preliminary and are presented to demonstrate the technique. Correspondence is invited by anyone working with a similar project who may either have suggestions or questions.

K. Steinbach
F. B. Varnum
U. S. Army Engr. Res. and Dev. Labs.
Fort Belvoir, Va.

## Maximum Gain in Monopulse Difference Mode*

It is well known that maximum theoretical gain of a simple antenna of the large aperture type is obtained if the aperture is uniformly illuminated with no spillover.[1] For a monopulse antenna, it has been shown that maximum theoretical slope gain in the difference mode is obtained with odd linear illumination having no spillover.[2]

It is generally accepted that slope gain is the most significant measure of performance in the difference mode, because the slope at the center of the difference pattern is a factor in the ability of a monopulse antenna to accurately track a target. However, the gain at the peaks of the difference pattern is also significant, since it may affect the ability of a monopulse antenna to achieve angle lock-on of a distant target. Furthermore, peak gain is a term which is universally employed in the antenna field, and is a standard measurement in the testing of an antenna. In this note the aperture illumination which provides maximum theoretical peak gain in the difference mode will be determined.

Consider first the field which a distant source would create on the aperture of an antenna during reception. If the source is located off the antenna axis, the distribution of this field may be separated into even and odd components. Since only the odd component couples to the difference mode of the antenna, only this component is of interest here. The shape of this odd field distribution is a simple sine curve, as indicated in Fig. 1.

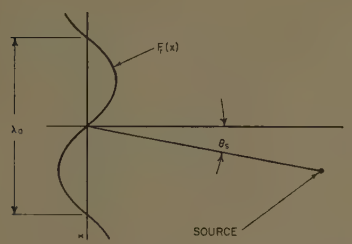



Fig. 1—Odd component of field received from a distant off-axis source.

It is described by the following relationships:

$$F_r(x) = c_1 \sin kx \tag{1}$$

$$k = \frac{2\pi}{\lambda_a} = \frac{2\pi \sin \theta_s}{\lambda} \tag{2}$$

where $F_r(x)$ is the amplitude of the received field distribution, $c_1$ is a constant, $x$ is the aperture coordinate, $\lambda_a$ is the wavelength of the sine distribution across the aperture, $\theta_s$ is the angle of the source, and $\lambda$ is the free-space wavelength.

What shape should the aperture illumination of this antenna (during transmission) have in order to obtain a maximum signal (during reception) from the above field distribution? By analogy with the "matched filter criterion"[3] the illumination should have exactly the same shape as the field distribution. Thus an aperture illumination having a sine shape should yield maximum gain the difference mode. The wavelength of the sine illumination should equal the wavelength of the field distribution, which, in turn, is directly related to the angle of the source. The following relationship describes this matching aperture illumination:

$$F_t(x) = c_2 \frac{\sin kx}{\left( \int_{-L/2}^{L/2} \sin^2 kx dx \right)^{1/2}} \tag{3}$$

where $F_t(x)$ is the amplitude of the aperture illumination during transmission, $C_2$ is a constant, and $L$ is the extent of a one-dimensional aperture. The denominator term conserves power across the aperture.

When the antenna aperture is finite in extent, there is one source angle which permits the greatest signal to be obtained. This angle is the angle of the peak of the difference pattern. Correspondingly, there is one wavelength of the sine illumination which yields the maximum peak gain with a given finite aperture. This may be determined in the following manner.

The voltage received by an antenna is given by

$$E = c_3 \int_{-L/2}^{L/2} F_t(x) F_r(x) dx \tag{4}$$

where $E$ is the received voltage and $c_3$ is a constant. This is easily demonstrable by means of superposition and reciprocity, or by analogy to filter theory. Now combining (1), (3), and (4), the received voltage is

$$E = c_4 \int_{-L/2}^{L/2} \frac{\sin^2 kx}{\left( \int_{-L/2}^{L/2} \sin^2 kx dx \right)^{1/2}} dx. \tag{5}$$

Evaluating this integral gives the following simple relation for received voltage as a function of $k$ and $L$:

$$E = \frac{c_4}{2} \left( L - \frac{\sin kL}{k} \right)^{1/2}. \tag{6}$$

In order to obtain the relation between $k$ and $L$ which gives the maximum received voltage, (6) is differentiated with respect to $k$ and the derivative is set equal to zero. This gives

$$\tan kL = kL. \tag{7}$$

Solving this transcendental equation gives

$$\frac{kL}{2} = 0.715\pi$$

$$= \frac{\pi L}{\lambda_a} = 128.7°. \tag{8}$$

Thus the wavelength of the sine illumination for maximum peak gain in the difference mode has been determined.

The resulting illumination is shown, normalized to its peak, in Fig. 2. This illumination is a truncated sine curve having an edge taper of 2.15 db, and happens to closely resemble that which exists in many monopulse antennas of the type in which a parabolic reflector is illuminated by a monopulse feed. However, in those cases the gain in the difference mode is far below its maximum possible value, because of the large amount of spillover. Naturally, there is no spillover for this theoretical case.

The antenna pattern may be calculated for this illumination by means of the usual techniques,[4] or may be inferred from the truncated-sine illumination. The pattern, normalized to its peak, is shown in Fig. 3.

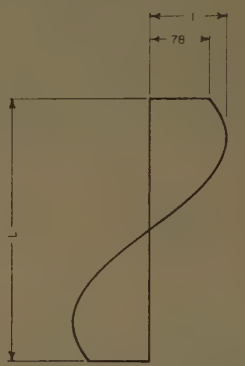



Fig. 2—Aperture illumination for maximum gain at the difference peak.

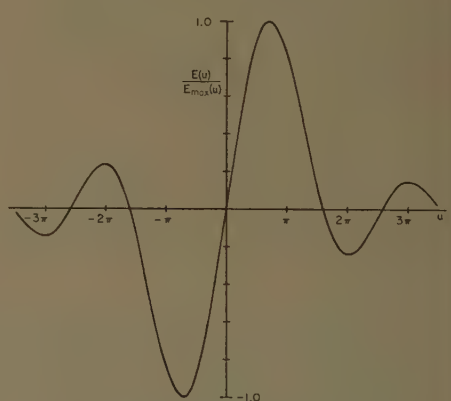



Fig. 3—Difference pattern having maximum peak gain.

Its value, normalized to the unit-amplitude illumination of Fig. 2, is given by the following formula:

$$E(u) = \frac{1}{2} \frac{\sin (u - u_m)}{u - u_m} - \frac{1}{2} \frac{\sin (u + u_m)}{u + u_m}$$

where

$$u = \frac{\pi L \sin \theta}{\lambda} \tag{9}$$

and

$$u_m = \frac{\pi L \sin \theta_m}{\lambda} = 0.715\pi$$

* Received by the PGAP, December 19, 1960.

[1] S. Silver, "Microwave Antenna Theory and Design," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co. Inc., New York, N. Y., vol. 12, pp. 177–178; 1949.

[2] G. M. Kirkpatrick, General Electric Co., ASTIA Doc. No. AD 18458, pp. 13–15; August 1952. Also see E. J. Powers, ASTIA Doc. No. AD 231217, pp. 8–15, 66–72; July, 1959.

[3] J. H. Van Vleck and D. Middleton, "A theoretical comparison of visual, aural, and meter reception of pulsed signals in the presence of noise," *J. Appl. Phys.*, vol. 17, pp. 943–944, 964; November, 1946.

[4] J. F. Ramsay, "Fourier transforms in aerial theory—Part II, Fourier sine transforms," *Marconi Rev.*, vol. 10, p. 17; January–March, 1947.

where $E(u)$ is the pattern amplitude, $\theta$ is the pattern angle, and $\theta_m$ is the angle of the pattern peak. This pattern is simply the difference of two ordinary maximum-gain patterns, squinted to the difference peak angles. It is interesting that the squint angle is such that the first sidelobe peak of one of these "lobe" patterns coincides with the main peak of the other; this condition would be expected to correspond to maximum peak gain in the difference pattern if two such lobe patterns could be independently created with a single antenna structure.

The value for difference peak gain may be expressed by comparing it with ordinary peak gain. This ratio is determined as follows:

$$\frac{g}{g_0} = \frac{E^2(u_m)}{\frac{1}{L}\int_{-L/2}^{L/2} \sin^2\left(0.715\pi \frac{2x}{L}\right) dx} \qquad (10)$$

where $g$ is the maximum difference peak gain, and $g_0$ is the ordinary maximum peak gain obtained with uniform illumination of the same aperture, both cases, of course, assuming no spillover. The denominator represents the power of the unit-amplitude illumination of Fig. 2. Evaluating this expression yields

$$\frac{g}{g_0} = \sin^2 0.715\pi = 0.609. \qquad (11)$$

Thus the maximum possible gain in the difference mode of a monopulse antenna is 2.15 db below the maximum possible gain in the sum mode of the same antenna. It is interesting that this number exactly equals the 2.15-db edge taper mentioned in connection with Fig. 2. It may also be noted that this gain ratio is reasonably close to the 3-db ratio which would be expected if the two previously-mentioned "lobe" components of the difference pattern were separated by a squint angle great enough so that only one component contributed significantly at the difference peak.

In conclusion, it is appropriate to mention that a knowledge of the illumination which provides maximum theoretical peak gain in the difference mode of a monopulse antenna can be helpful in the design of such an antenna, even if one chooses to avoid this illumination because of the high sidelobes it yields. Furthermore, the value of the theoretical peak gain can provide a useful standard for rating the efficiency of a monopulse antenna. Of course, the results presented here apply exactly only to a one-dimensional aperture or to a rectangular aperture with separable distributions.

It is perhaps also worth noting that the "matching" method employed to determine the optimum illumination involved a simple physical process rather than a complex mathematical one. The method is applicable to other similar problems; for example, the illumination for ordinary maximum peak gain and for maximum difference slope gain, mentioned at the beginning of this communication, may be solved by inspection with this method.

PETER W. HANNAN
Wheeler Labs.
Smithtown, N. Y.

## Current Distributions on Cylinders Excited by Spherical Electromagnetic Waves*

### INTRODUCTION

Relationships are established between the current distributions on infinitely-long perfectly-conducting cylindrical scatterers of arbitrary cross section, illuminated by 1) plane waves and 2) spherical waves. These are of value because, although most analytical treatments of scattering by cylinders assume plane-wave excitation, most physical situations involve spherical- or cylindrical-wave illumination.

It is known that the cross-sectional distributions of current on circular and elliptic cylinders[1-3] excited by a spherical wave with a far-zone phase center are only locally dependent on the incident field. An important example of this is the following: Suppose that the incident spherical wave at some cross section of a circular or elliptic cylinder is identical locally to a plane wave propagating at some angle to the cylinder axis; then, the transverse-current distributions at that cross section for both illuminations are also identical. This result is valid for cylinders of more general cross section, subject to certain restrictions on cylinder size and source distance, and is proven for waves due to electric sources polarized in a plane perpendicular to any radial line [see $\rho_1$ of Fig. 1(b)] and, specifically, for sources polarized parallel and perpendicular to the cylinder axis in that plane. Vector addition recovers the more general case.

An additional relationship between the transverse-current distributions at different parallel cross sections of a cylinder when excited by a spherical wave, or at any cross section when excited by plane waves incident at different angles—a relationship that is known to hold for circular and elliptic cylinders[1-3]—is shown to extend to arbitrary cylinders: The components of the transverse current distributions for $\theta_0 = \pi/2$ are identical to those same components at any other value of $\theta_0$ on a larger cylinder—larger by a cross-sectional scale factor $(\sin \theta_0)^{-1}$. [See Fig. 1(a) and (b).] Therefore, the functional dependence of these components depends upon the product of $\sin \theta_0$ and a characteristic dimension.

With these relations, much of the data on current distribution induced on a cylinder by a spherical or an oblique plane wave can be obtained from the normal plane-wave case. Experimental simplification is also suggested, since current distributions for normal plane-wave incidence on a range of cylinders of the same cross section but of various sizes can be obtained through the use of one cylinder of the largest size desired, and a source of spherical waves. The current distributions on all smaller cylinders may then be obtained by varying, axially, the cross section at which transverse-current measurements are made.

Proof of these properties utilizes integral equations for the electromagnetic field[4] recast into an appropriate form by a Fourier integral operation, and integrated by the method of steepest descents to obtain the first terms of series expansions in powers of $(kR_0)^{-1}$. Retaining only the first terms provides results valid in the far zone of the spherical-wave source, $kR_0 \gg 1$.

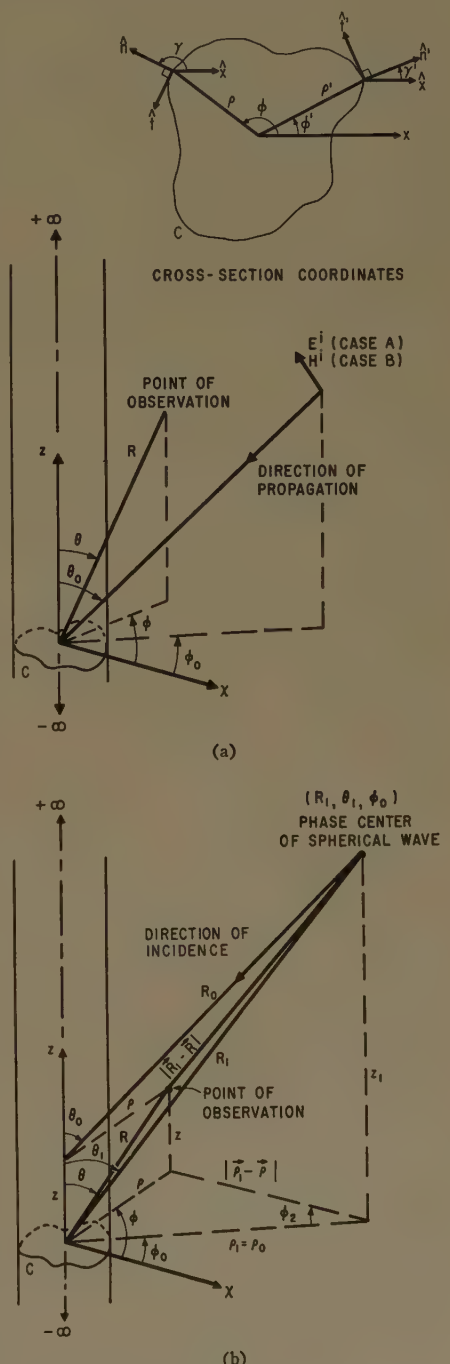



Fig. 1—(a) Configuration for plane-wave scattering; (b) configuration for spherical-wave scattering.

* Received by the PGAP, January 6, 1961. This research was made possible by support extended by Cruft Lab., Harvard University, Cambridge, Mass., under contract No. AF19(604)-786.

[1] W. S. Lucke, "Electric dipoles in the presence of elliptic and circular cylinders," *J. Appl. Phys.*, vol. 22, pp. 14–19; January, 1951.

[2] G. Sinclair, "The patterns of antennas located near cylinders of elliptical cross section," PROC. IRE, vol. 39, pp. 660–668; June, 1951.

[3] P. S. Carter, "Antenna arrays around cylinders," PROC. IRE, vol. 31, pp. 671–693; December, 1943.

[4] J. A. Stratton, "Electromagnetic Theory," McGraw-Hill Book Company, Inc., New York, N. Y., p. 466; 1941.

## Formulation of the Integral Equations

The integral equations are:[5]

*Case A*

$\boldsymbol{H}^i$, the incident magnetic field, perpendicular to cylinder axis $z$:

$$E_z^i(S) = -i\omega\mu \int_{-\infty}^{\infty} dz' \left\{ \oint_{c'} \left(1 + \frac{1}{k^2}\frac{\partial^2}{\partial z^2}\right) G(S,S')K_z(S')dc' \right\}; \quad (1a)$$

*Case B*

$\boldsymbol{E}^i$, the incident electric field, perpendicular to cylinder axis, $z$:

$$H_t^i(S) = K_z(S) - \oint_{-\infty}^{\infty} dz' \left\{ \int_{c'} K_{t'}(S') \sin(\gamma' - \gamma) \frac{\partial G(S',S)}{\partial z'} + K_z(S') \left[ \cos(\gamma - \phi') \frac{\partial}{\partial \phi'} - \frac{\sin(\phi' - \gamma)}{\rho} \frac{\partial}{\partial \phi'} \right] G(S',S)dc' \right\} \quad (1b)$$

$$H_z^i(S) = -K_t(S) - \oint_{-\infty}^{\infty} dz' \left\{ \int_{c'} K_{t'}(S) \left[ \frac{\sin(\phi' - \gamma')}{\rho'} \frac{\partial}{\partial \phi'} - \cos(\phi' - \gamma') \frac{\partial}{\partial \rho'} \right] G(S',S)dc' \right\}, \quad (1c)$$

where coordinates are as defined in Fig. 1(a), and where either $\boldsymbol{E}^i$ or $\boldsymbol{H}^i$ is perpendicular to the cylinder axis; where $c$ is cylinder cross-section surface, $S$ is shorthand for position coordinate on $c$, $\boldsymbol{K}(S) = \hat{n} \times \boldsymbol{H}(S)$ = surface-current density, $\hat{n}$ is the outward normal from $c$, and

$$G(R, R') = \exp ik|\boldsymbol{R} - \boldsymbol{R}'|/4\pi|\boldsymbol{R} - \boldsymbol{R}'|.$$

## Plane-Wave Incidence [Fig. 1(a)]

*For Case A:*

$$E_z^i = A(\rho, \phi, z; \theta_0, \phi_0) \sin\theta_0. \quad (2a)$$

*For Case B:*

$$H_t^i = -A(\rho, \phi, z; \theta_0, \phi_0) \sin(\theta_0 - \gamma) \cos\theta_0 \quad (2b)$$

$$H_z^i = A(\rho, \phi, z; \theta_0, \phi_0) \sin\theta_0. \quad (2c)$$

$A \equiv \exp -ik[z\cos\theta_0 + \rho\sin\theta_0\cos(\phi - \phi_0)]$; $\theta_0$, $\phi_0$ indicate the incident wave's propagating direction.

Utilizing the Fourier integral form of $G(\boldsymbol{R}, \boldsymbol{R}')$[6]—the fact that $\boldsymbol{K}(S)$ can be represented as $\mathcal{K}(s) \exp -ikz\cos\theta_0$—and the identity

$$2\pi\delta(x) = \int_{-\infty}^{\infty} \exp ixz'dz'$$

in the right sides of (1a) through (1c), and (2a) through (2c) on the left sides, we have the following result ($s$ is the two-dimensional position coordinate on $c$):

*Case A*

$$\frac{A(\rho, \phi, z; \theta_0, \phi_0)}{\sin\theta_0} = \frac{\omega u}{4} \oint_{c'} H_0^{(1)}(|\varrho - \varrho'| k \sin\theta_0) K_z(\rho', \phi', z)dc'. \quad (3a)$$

*Case B*

$$-A(\rho, \phi, z; \theta_0, \phi_0) \sin(\phi_0 - \gamma)\cos\theta_0 = K_z(\rho, \phi, z) - \frac{i}{4} \oint_{c'} dc' \left\{ ik\cos\theta_0 K_{t'}(\rho', \phi', z) \sin(\gamma' - \gamma) H_0^{(1)}(|\varrho - \varrho'| k\sin\theta_0) + K_z(\rho', \phi', z) \left[ \cos(\gamma - \phi') \frac{\partial}{\partial \rho'} - \frac{\sin(\phi' - \gamma)}{\rho'} \frac{\partial}{\partial \phi'} \right] H_0^{(1)}(|\varrho - \varrho'| k\sin\theta_0) \right\} \quad (3b)$$

$$A(\rho, \phi, z; \theta_0, \phi)\sin\theta_0 = -K_t(\rho, \phi, z) - \frac{i}{4} \oint_{c'} K_{t'}(\rho', \phi', z) \left[ \frac{\sin(\phi' - \gamma)}{\rho'} \frac{\partial}{\partial \phi'} - \cos(\phi' - \gamma') \frac{\partial}{\partial \rho} \right] H_0^{(1)}(|\varrho - \varrho'| k\sin\theta_0) dc'. \quad (3c)$$

In each case, $\rho$ and $\rho'$ are on $c$ and $c'$. Note that the cross-sectional dependence of $K_z$, case A, and $K_t$, case B, are identical to those on a cylinder of identical shape illuminated by a normally-incident plane wave but with transverse dimension scaled-down by $\sin\theta_0$. Note also that for normal incidence $K_z$, case B vanishes.

## Spherical Incident Waves [Fig. 1(b)]

The sources are polarized in the $z$ direction, case A, or perpendicular to the $\rho_1 - z_1$ plane, case B, at $R_1$, $\theta_1$, $\phi_0$. The incident fields are representable as follows:

*Case A*

$$E_z^i = B \int_{-\infty}^{\infty} F_a(k, \alpha, \phi_2) H_0^{(1)}(|\varrho - \varrho_1| \sqrt{k^2 - \alpha^2}) e^{i\alpha(z - z')} d\alpha; \quad (4a)$$

*Case B*

$$H_t^i = \omega\epsilon\mu C \sin(\phi_1 - \gamma) \int_{-\infty}^{\infty} \alpha F_b(k, \alpha, \phi_2) H_0^{(1)}(|\varrho - \varrho_1| \sqrt{k^2 - \alpha^2}) e^{i\alpha(z - z_1)} d\alpha \quad (4b)$$

$$H_z^i = i\omega\epsilon\mu C \int_{-\infty}^{\infty} F_b(k, \alpha, \phi_2) \frac{\partial}{\partial \rho_1} H_0^{(1)}(|\varrho - \varrho_1| \sqrt{k^2 - \alpha^2}) e^{i\alpha(z - z_1)} d\alpha; \quad (4c)$$

where the $F$'s are field factors, assumed to be well-behaved in $\alpha$ and $\phi_2$ with possible singularities in the $\alpha$-plane only at points $\alpha = \pm k$, the branch points of $H_0^{(1)}$.

Substituting (4a) through (4c) for the left sides of (1a) through (1c), and utilizing the Fourier integral of $\boldsymbol{K}$,

$$\boldsymbol{K}(\rho_1, \phi_0, \phi_2, z_1, S') = \int_{-\infty}^{\infty} \mathcal{K}(\rho_1, \phi_0, \phi_2, z_1, s', \beta) e^{i\beta s'} d\beta$$

(the $\rho_1$, $\phi_0$, $\phi_2$ and $z_1$ factors are introduced to exhibit the functional dependence on source position)—and the aforementioned $\delta$-function identity in the right sides, equations are obtained in which both sides are of the form of Fourier integrals. We shall refer to them as (5a)–(5c) despite the fact that space limitations prevent their explicit presentation. Since the vector components of the incident fields are, at worst, sectionally continuous on the surface, the integrands (or transforms) are identical (in the sense of Fourier convergence).

In the limit $k|\varrho_1 - \varrho| \gg 1$, $\rho_1 \gg \rho$, with $\rho$ on $c$, and the cylinder in the far zone of the source, (4a) through (4c) [≡ the left sides of (5a) through (5c)] are integrable by the method of steepest descents.[7] First-order integrations valid in the above limit yield terms of order $(kR_0)^{-1}$.[8] These represent the first terms in series of inverse powers of $kR_0$ (in which higher terms are negligible), and hence are equal to like terms on the other sides [(5a) through (5c)].

Equality of integrands (in the Fourier sense) on each side of (5a) through (5c) specifies also equality of form and hence an identical far-zone relationship on the right of (5a) through (5c) between each integral and its integrand.[8] Further examination of (5) shows that $\mathcal{K}$ may be decomposed in the limit $\rho_1 \gg \rho$, $k|\varrho_1 - \varrho| \gg 1$, since

$$\mathcal{K}(\rho_1, \phi_0, \phi_2, z_1, s, \alpha) = h(\rho_1, \phi_0, \phi_2, s, \alpha) \exp i[\rho_1\sqrt{k^2 - \alpha^2} - \alpha z_1]$$

where $h$ is a slowly-varying function of $\alpha$.

Utilizing the above relationship and the Fourier integral, the first term in the series of powers of $(kR_0)^{-1}$ for $\boldsymbol{K}$ is, in the limit,

$$\mathcal{K}(\rho_1, \phi_0, \phi_2, z_1, S) = \sqrt{\frac{2\pi}{kR_0}}\, \mathcal{K}(\rho_1, \phi_0, \phi_2, z, s, k\cos\theta_0) k\sin\theta_0 \cdot \exp -i\left(\frac{\pi}{4} - kz\cos\theta_0\right).$$

[5] These may be obtained by applying the boundary conditions for a perfectly conducting surface to the equations given by Stratton,[4] then specializing them to the appropriate coordinate system and incident-field polarization. Rationalized mks units and time dependence $e^{-i\omega t}$ are assumed.

[6] $G(R, R') = \frac{i}{8\pi} \int_{-\infty}^{\infty} H_0^{(1)}(|\varrho - \varrho'| \sqrt{k^2 - \alpha^2}) \exp i\alpha(z - z')d\alpha.$

[7] The integrands are of the form $D(\rho_1, \phi_0, \phi_2, \rho, \phi, k, \alpha) \exp i[\alpha(z - z_1) + \rho_1\sqrt{k^2 - \alpha^2}]$, where $D$ is a slowly-varying function of $\alpha$ compared to the exponential. The saddle points are at $\alpha = k\cos\theta_0$.

[8] Each term is $k\sqrt{2\pi/kR_0}\sin\theta_0 \exp -i\pi/4$ times its integrand, evaluated at the saddle point, $\alpha = \cos\theta_0$.

Inserting the indicated saddle-point evaluations on both sides of (5) and then substituting the above representation for $\mathcal{K}$, the following are obtained, in the limit, for terms of order $(kR_0)^{-1}$: (It is important to note that the $(kR_0)^{-1}$ dependences on the right-hand sides of these equations are carried implicitly by the $K$'s.)

*Case A*

$$\frac{2ikBF_a(k, k\cos\theta_0, \phi_2)}{kR_0\sin^2\theta_0}\exp ik[R_0 - \rho\sin\theta_0\cos(\phi-\phi_0)] = \{\text{right side of (3a)}\}; \quad (6a)$$

*Case B*

$$\frac{2ik^4}{\omega kR_0}C\sin(\phi_0-\gamma)\cos\theta_0 F_b(k, k\cos\theta_0, \phi_2)\exp ik[R_0-\rho\sin\theta_0\cos(\phi-\phi_0)] = \{\text{right side of (3b)}\} \quad (6b)$$

$$-\frac{2ik^4}{\omega kR_0}CF_b(k, k\cos\theta_0, \phi_2)\sin\theta_0\exp ik[R_0 - \rho\sin\theta_0\cos(\phi-\phi_0)] = \{\text{right side of (3c)}\}; \quad (6c)$$

each evaluated in the limit for $\rho$ on $c$, $\rho'$ on $c'$ and with $K(\rho_1, \phi_0, \phi_2, z_1, s, z)$ replacing $K(\rho', \phi, z)$.

These are the desired forms of the $(kR_0)^{-1}$ terms of the integral equation for spherical-wave illumination. If the ratio of the largest dimension of the cylinder cross section to the source distance is large, and if the far-zone relationship is valid (the limiting conditions assumed above), $F_a$ and $F_b$ are constants in $\phi_2$ on the cylinder surface. Then, for equal values of $\theta_0$ (polar angle of incidence), the cross-sectional dependences of the left sides of (6) and (3), respectively, are equal, and hence the respective cross-sectional distributions of current are identical to within a multiplicative constant (dependent upon source strength, distance, and axial position). As before, the cross-sectional dependences of the components of current existing on a cylinder scaled-down under normal illumination by sin $\theta_0$ are equal to those on the unreduced cylinder at the cross section where the angle of incidence is $\theta_0$.

D. B. Brick
Applied Research Lab.
Sylvania Electric Products, Inc.
Waltham, Mass.

## The Potential Utility of Scanning Microwave Beams in Plasma Diagnostics*

Summary

A technique is proposed for improving the data rate in certain plasma experiments wherein the electron density is measured. A scanning microwave beam is used in conjunction with a dc magnetic field to create a plasma resonance dependent on scan angle. Electron density in the plasma can then be measured by noting the beam angle where resonance occurs.

* Received by the PGAP, January 9, 1961. This study was supported by Lockheed Missiles and Space Division.

### Introduction

In the ever-growing study of plasmas, perhaps the most-sought property is the electron density. The interaction between an electromagnetic wave and the plasma has provided a useful tool for measuring this density. It is well-known[1] that in the absence of a dc magnetic field, the plasma can exhibit an equivalent dielectric constant given by

$$\epsilon = \epsilon_0\left(1 - \frac{\omega_p^2}{\omega^2}\right)$$

in which $\epsilon_0$ is the free-space permittivity, $\omega$ is the angular frequency of the electromagnetic wave, and

$$\omega_p^2 = \frac{ne^2}{m\epsilon_0} \quad (2)$$

with $n$ the electron density, $e$ the magnitude of the charge on the electron, and $m$ its mass. $\omega_p$ is called the plasma frequency. A plane wave propagating through the plasma will experience a resonance, when $\omega_p = \omega$, which is clearly discernible in either the transmitted wave or the reflected wave. Hence, for a plasma with fixed properties, varying $\omega$ till resonance is observed gives a measure of $\omega_p$ and thus $n$. Similarly, for a plasma of varying properties, use of a judiciously chosen fixed $\omega$ will give knowledge of $\omega_p$ and $n$ at the instant resonance is observed. Both experiments form a part of plasma diagnostics, but the second experiment is far more common.

A major difficulty in the technique just described is that, unless the boundaries are well defined, only one piece of data at resonance is obtained for each performance of the experiment. In many situations, it is desired to know $\omega_p$ as a function of time. For a repeatable experiment, $\omega_p(t)$ can be synthesized by performing the measurement over and over again for a sequence of $\omega$ values. However, this can be costly and introduces additional sources of error.

By using a dc magnetic field perpendicular to both the electric and Poynting vectors of the electromagnetic wave, a triple resonance phenomenon can be created[2] offering the possibility of an increased data rate. However, the increase is modest, and clearly what is preferred is a technique for measuring the entire curve $\omega_p(t)$ in one experiment.

Accordingly, a search has been undertaken for such a technique, with some small measure of success, as will be shown in what follows.

### Analysis

In a plasma region pervaded by a uniform dc magnetic field, which for convenience is taken to be $z$-oriented, it can be shown[3] that the equivalent permittivity is a tensor, given by expression

$$\epsilon = \epsilon_0\begin{bmatrix}\epsilon_{11} & \epsilon_{12} & 0\\ -\epsilon_{12} & \epsilon_{11} & 0\\ 0 & 0 & \epsilon_{33}\end{bmatrix} \quad (3)$$

in which $\epsilon_0$ is the permittivity of free space and

$$\epsilon_{11} = 1 + \frac{\omega_p^2}{j\omega}\frac{\nu + j\omega}{(\nu+j\omega)^2 + \omega_b^2} \quad (4)$$

$$\epsilon_{12} = -\frac{\omega_p^2}{j\omega}\frac{\omega_b}{(\nu+j\omega)^2+\omega_b^2} \quad (5)$$

$$\epsilon_{33} = 1 + \frac{\omega_p^2}{j\omega}\frac{1}{(\nu+j\omega)}. \quad (6)$$

In the foregoing, $\nu$ is the collision frequency, and

$$\omega_b = \frac{eB_0}{m} \quad (7)$$

is the cyclotron frequency, with $B_0$ the strength of the dc magnetic field.

Assume a plane wave is traversing the plasma region with its Poynting vector making an angle $\theta$ with respect to the dc magnetic field. No loss in generality results from placing the Poynting vector in the $x$-$z$ plane. Then each component of the plane wave is expressible as a constant times the exponential factor

$$e^{j(\omega t - k_x X - k_z Z)} \quad (8)$$

with

$$k_x = k\sin\theta \quad (9)$$

$$k_z = k\cos\theta \quad (10)$$

in which $k$ is the complex propagation constant of the plane wave.

Using (8), the tensor wave equation

$$\nabla\times\nabla\times\boldsymbol{E} = \omega^2\mu_0\epsilon\boldsymbol{E} \quad (11)$$

yields the three component relations

$$\begin{cases}(k_0^2\epsilon_{11} - k_z^2)E_x + k_0^2\epsilon_{12}E_y + k_xk_zE_z = 0\\ k_0^2\epsilon_{12}E_x + (k^2 - k_0^2\epsilon_{11})E_y = 0\\ k_xk_zE_x + (k_0^2\epsilon_{33} - k_x^2)E_z = 0.\end{cases} \quad (12)$$

To obtain a nontrivial solution of (12), one sets the determinant formed from the coefficients of the components of $\boldsymbol{E}$ equal to zero. When this is done and use is made of (9) and (10), we obtain the dimensionless expression

$$[\epsilon_{11}\sin^2\theta + \epsilon_{33}\cos^2\theta]\left(\frac{k}{k_0}\right)^4 - [(\epsilon_{11}^2)+\epsilon_{12}^2)\sin^2\theta + \epsilon_{11}\epsilon_{33}(1+\cos^2\theta)]\left(\frac{k}{k_0}\right)^2 + (\epsilon_{11}^2 + \epsilon_{12}^2)\epsilon_{33} = 0. \quad (13)$$

[1] See, e.g., S. K. Mitra, "The Upper Atmosphere," The Asiatic Society, Calcutta, India, 2nd ed., ch. 6; 1952.
[2] *Ibid.*, Fig. 9.
[3] W. P. Allis, "Motion of ions and electrons," in "Handbuch der Physik," vol. 21, pp. 383–444; 1956.

It is instructive to examine (13) for the special case $\theta=\pi/2$. The solutions are then easily shown to be

$$\left(\frac{k}{k_0}\right)^2 = \frac{\epsilon_{11}^2+\epsilon_{12}^2+\epsilon_{11}\epsilon_{33} \pm [\epsilon_{11}^2+\epsilon_{12}^2-\epsilon_{11}\epsilon_{33}]}{2\epsilon_{11}}. \quad (14)$$

We shall refer to the minus sign solution as case 1 and the plus sign solution as case 2. inserting (14) back into (12) and noting that $k_z=0$ for both cases (since $\theta=\pi/2$) we find, neglecting loss:

*Case 1*

$$\left(\frac{k}{k_0}\right)^2 = \epsilon_{33} = 1-\frac{\omega_p^2}{\omega^2}, \quad (15)$$

$$E_x = E_y = 0 \quad E_z \neq 0. \quad (16)$$

This is the conventional case referred to in the introduction and is independent of the presence of the magnetic field. One resonance occurs when $\omega_p=\omega$.

*Case 2*

$$\left(\frac{k}{k_0}\right)^2 = \frac{\epsilon_{11}^2+\epsilon_{12}^2}{\epsilon_{12}} = \frac{\left(\omega-\omega_b-\frac{\omega_p^2}{\omega}\right)\left(\omega+\omega_b-\frac{\omega_p^2}{\omega}\right)}{\omega^2-\omega_b^2-\omega_p^2}, \quad (17)$$

$$E_x = -\frac{\epsilon_{12}}{\epsilon_{11}}E_y \neq 0 \quad E_z = 0. \quad (18)$$

This is the second case referred to in the introduction and provides three resonances, when

$$\omega_p^2 = \omega(\omega-\omega_b), \quad (19)$$

$$\omega_p^2 = (\omega+\omega_b)(\omega-\omega_b), \quad (20)$$

$$\omega_p^2 = \omega(\omega+\omega_b). \quad (21)$$

In principle, these three resonances can be spread apart at will merely by choosing $\omega_b$ to be a big enough fraction of $\omega$. In practice, it is difficult to get large fractions if $\omega$ is high up in the microwave range because of the large values of $B_0$ required. However, even as high as $X$ band, $\omega_b$ can be made large enough to separate the three resonances sufficiently to argue that a triple data rate is possible.

We therefore draw the important conclusion that the two branch solutions of (13) have, respectively, one and three resonances as $\theta\rightarrow\pi/2$. It seems reasonable to expect that these resonances will still be evident at angles of $\theta$ other than $\pi/2$. Since the coefficients in (13) are functions of $\theta$, $\omega_p$ at resonance should be a function of $\theta$ as well as a function of $\omega$ and $\omega_b$. This opens up the possibility that a *scanning* microwave beam ($\theta$ variable) will yield a resonant relation $\omega_p(\theta)$ for fixed values of $\omega$ and $\omega_b$. Thus, in experiments in which $\omega_p$ is varying, recording the variable angle $\theta$ at which resonance is observed might give data which could be translated into $\omega_p(t)$.

The branch of case 1 is initially more attractive, since a triple resonance is not needed if sufficient dynamic range is available in the resonance relation $\omega_p(\theta)$ for the single resonance case. It is convenient to make the substitutions

$$a = \epsilon_{11}, \quad (22)$$

$$b^2 = \epsilon_{11}\epsilon_{33}, \quad (23)$$

$$c^2 = \epsilon_{11}^2+\epsilon_{12}^2, \quad (24)$$

$$\zeta = \left(\frac{k}{k_0}\right)^2, \quad (25)$$

which convert (13) to

$$[a^2\sin^2\theta+b^2\cos^2\theta]\zeta^2 - [ac^2\sin^2\theta+ab^2(1+\cos^2\theta)]\zeta + b^2c^2 = A\zeta^2+B\zeta+C=0, \quad (26)$$

the general solution of which is, of course,

$$\zeta = -\frac{B}{2A} \pm \sqrt{\left(\frac{B}{2A}\right)^2-\frac{C}{A}}. \quad (27)$$

From (15), we note that the resonance in case 1 is a zero. Likewise, from (17), it is clear that the resonances defined by (19) and (21) are zeros. The motions of all three of these zeros with $\theta$ can be deduced by setting $\zeta=0$ in (27). But this results in the condition

$$\frac{C}{A}=0=\frac{\epsilon_{33}(\epsilon_{11}^2+\epsilon_{12}^2)}{\epsilon_{11}\sin^2\theta+\epsilon_{33}\cos^2\theta}. \quad (28)$$

We assume that $B_0\neq 0$ and that $\omega_p$ is finite. Then $\epsilon_{11}\neq\epsilon_{33}$, and in the range $0<\theta<\pi/2$, the denominator is never zero nor infinite. Either

$$\epsilon_{33}=0 \quad (29)$$

or

$$\epsilon_{11}^2+\epsilon_{12}^2=0. \quad (30)$$

But (29) is equivalent to $\omega_p=\omega$, the zero resonance of case 1, and (30) is equivalent to (19) and (21), the zero resonances of case 2. Thus none of the three zero resonances move with $\theta$. case 1 is ruled out as a possibility, and in case 2, the only hope is the pole described by (20).

To study the behavior of the pole, we set $1/\xi=0$, which leads to the condition

$$\frac{A}{C}=0=\frac{\epsilon_{11}\sin^2\theta+\epsilon_{33}\cos^2\theta}{\epsilon_{33}(\epsilon_{11}^2+\epsilon_{12}^2)}, \quad (31)$$

which has as its principal result

$$\epsilon_{11}\sin^2\theta+\epsilon_{33}\cos^2\theta=0 \quad 0<\theta\leq\frac{\pi}{2}. \quad (32)$$

Using (4) and (6) and neglecting loss, (32) can be put in the form

$$\frac{\omega_p^2}{\omega^2}=\frac{1-\frac{\omega_b^2}{\omega^2}}{1-\frac{\omega_b^2}{\omega^2}\cos^2\theta}, \quad (33)$$

and the plasma frequency at pole resonance is seen to depend on the angle $\theta$. A plot of (33) is shown in Fig. 1.

Inspection of this figure reveals that if the cyclotron frequency $\omega_b$ is a small fraction of the plane wave frequency $\omega$, only a small variation of plasma density required for resonance with scan angle occurs. It takes a precisely controlled small excess of $\omega$ over $\omega_b$ to give a big dynamic range in $\omega_p$ at resonance vs $\theta$. It is highly improbable that this precise a control can ever be achieved.

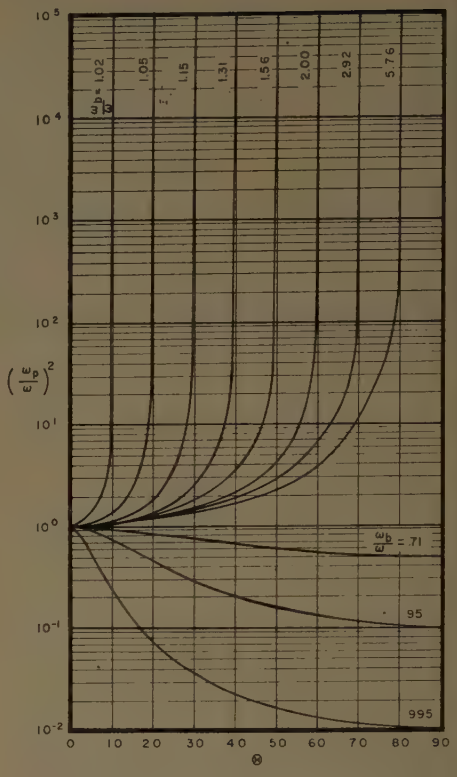



Fig. 1—Plasma density at resonance vs scan angle with dc magnetic field as parameter.

The situation is much more interesting if $\omega_b$ exceeds $\omega$. It is evident that several orders of magnitude of $\omega_p$ can be covered in a reasonable range of scan angles if $\omega_b$ is of the order of twice $\omega$, with the exact ratio not too critical.

This situation is not without its difficulties. $B_0$ should not be too large, as it might seriously affect the decay process in the plasma. This places an upper limit on $\omega$, which raises the problem of achieving a suitable beamwidth with the available aperture space. Thus, compromises must be made, but it appears that the technique should be practical in certain experiments and it seems well worth the try. The present state of the art insures reading the scan angle with sufficient accuracy to make use of the steep slope of the upper curves of Fig. 1. An extremely precise reading of scan angle might permit the measurement of plasma densities which are beyond present techniques.

## Application

The type of experiment in which this technique might be useful is suggested by Fig. 2. The transmitting and receiving antennas are electronically phased arrays. The phasing is synchronized so that the transmitting and receiving patterns always point in a common direction. If the phasing is cycled at a high rate compared to the plasma decay, the received signal can record the progress of the resonance with scan angle as the plasma density diminishes. Alternately, individual detectors can be placed on each element of the receiving array. Angular sensitivity can be increased by the use of conventional sum-and-difference techniques.

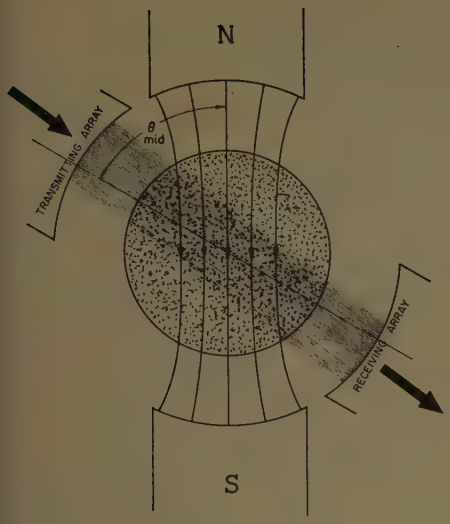



Fig. 2—Possible experimental setup.

### Conclusions

Four resonances are possible in the interaction of an electromagnetic wave and a plasma in the presence of a dc magnetic field. Three of these are zeros and are unaffected by the angle between the Poynting vector and the dc magnetic field. The fourth is a pole and is angle sensitive. For values of the cyclotron frequency above the plane-wave frequency, big dynamic ranges occur in the resonant value of plasma density as the angle is changed. Thus, a scanning beam offers the possibility of becoming a useful diagnostic tool.

### Acknowledgment

The author wishes to express his thanks to the staff of the Electromagnetics Laboratory, Lockheed Missiles and Space Division, for their counsel and support.

R. S. Elliott
Dept. of Engrg.
University of California
Los Angeles, Calif.
(Consultant to Lockheed Corp.)

## The Zone Plate as a Radio-Frequency Focusing Element*

### Introduction

More than a hundred years ago, Fresnel invented the optical zone plate, an interference device which has focusing properties similar to those of a lens. Application of the device seems to have been limited, and it has remained a laboratory curiosity. In the course of design studies on large parabolic reflectors for use in satellite communication, investigation of the zone plate was prompted by the apparent simplification it would offer in construction, in that it can be constructed on a plane surface. This note describes some of the properties of zone plates as potential substitutes for large paraboloids.

* Received by the PGAP, January 20, 1961.

### Basic Theory

The zone plate consists of a series of concentric opaque and transparent rings.

According to Huygen's principle, each point of a transparent zone acts as a secondary source which emits a spherical wave. Let the point $P$ be the desired focal point. Then, if the zones are of such a diameter that radiation from all the transparent zones arrive in phase, there will be constructive interference. This requires that the distance from $P$ to the edges of the successive zones be integral multiples of a half-wavelength. That is,

$$\rho_N = F + \frac{N\lambda}{2}, \tag{1}$$

where $\rho_N$ is the distance from the focal point to the $N$th zone edge, $F$ is the focal length, and $\lambda$ is the wavelength. Therefore, the radius of the $N$th edge is

$$r_N = \sqrt{\rho_N^2 - F^2}, \tag{2}$$

or

$$r_N = \frac{N\lambda}{2}\sqrt{1 + \frac{4F}{N\lambda}}. \tag{3}$$

Eq. (3) may be used for calculating the radii of the zones. However, we have found it usually simpler and quicker to use a graphical layout.

It is interesting to note that the focusing action is indifferent to whether the first zone is opaque or transparent. That is, the actions of complementary zone plates are identical.

This means that a given zone plate has a focus for reflection as well as transmission, since the opaque rings can be considered as reflection sources if they are conducting. If two feed antennas are located at these foci, and are properly phased, it should be possible to utilize both foci.

### Theoretical Gain

We now attempt to compute the theoretical gain of the zone plate. The amplitude $A_m$, due to the $m$th zone, is proportional to the area of the zone and inversely proportional to the distance from the zone to the focal point. Also, we must include an "obliquity factor" in Huygen's principle. The resulting equation is

$$A_m = K\frac{a_m}{d_m}(1 + \cos\theta), \tag{4}$$

where $a_m$ is the area of the $m$th zone, $d_m$ is the average distance from the zone to the focal point, $(1 + \cos\theta)$ is the obliquity factor and $K$ is a proportionality factor to be determined.

It requires only a bit of algebraic manipulation to derive the theoretical gain formulas.

Expressed in decibels, the gain for the aperture center zone plate is

$$G_{M,ap} = 20\log\left\{\frac{4\frac{F}{\lambda}+1}{4\frac{F}{\lambda}+\frac{1}{2}}\sum_{m=1}^{M}\frac{4\frac{F}{\lambda}+2m-\frac{3}{2}}{2\frac{F}{\lambda}+2m-\frac{3}{2}}\right\}, \tag{5}$$

and for the obstacle center zone plate, the gain is

$$G_{M,\mathrm{obs}} = 20\log\left\{\frac{\left(4\frac{F}{\lambda}+1\right)}{\left(4\frac{F}{\lambda}+\frac{1}{2}\right)}\sum_{m=1}^{M}\frac{\left(4\frac{F}{\lambda}+2m-\frac{1}{2}\right)}{\left(2\frac{F}{\lambda}+2m-\frac{1}{2}\right)}\right\}, \tag{6}$$

where $M$ is the number of zones in the plate.

Strictly speaking, (5) and (6) are correct only for zone plates imbedded in the center of an infinite screen. However, for a reasonably large number of rings, the obliquity factor comes into play, so that the outer zones become less and less important.

It is interesting to compare these theoretical gains with those of a paraboloid. As a specific example, the gain of a zone plate with $F/\lambda = 10$ has been calculated as a function of the number of zones. The corresponding diameters were then computed. The results are plotted in Fig. 1.

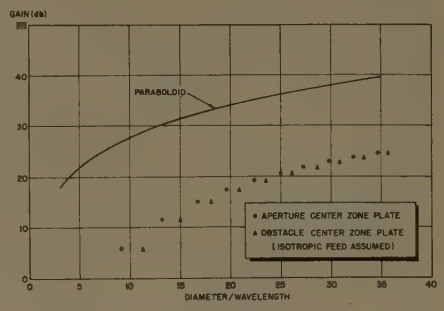



Fig. 1—Comparison of theoretical gains for zone plates and paraboloids.

### Use of a Ground Plane

As mentioned earlier, any zone plate has two principal foci, one for transmission and one for reflection. It therefore seems natural to attempt to fold the system with a ground plane and bring the two foci into coincidence.

The radiation that is reflected from the opaque zones will have undergone a 180° phase change. The radiation that passes through the transparent zones travels a distance $d$ to the ground plane, undergoes a phase reversal upon reflection, and returns through the distance $d$. In order for the radiation from both portions to arrive at the focal point in phase, the two portions should be 180° out of phase at the plane of the zone plate. Therefore, the spacing $d$ should be a quarter wavelength.

### Experimental Procedure

Experimental zone plates were built for the *X* band by cementing aluminum foil to plexiglass sheets, then laying out the rings with a beam compass, cutting along the lines with a razor blade, and finally peeling off the unwanted foil. Fig. 2 shows such a plate mounted on the antenna range turntable.

Fig. 2—Zone plate with ground plane mounted on antenna range turntable.

Two types of feed were tried—a straight dipole above a ground plane, and a turnstile in which the arms were drooped toward the ground plane in an effort to somewhat widen the pattern of the feed.

Fig. 3 is an example of the pattern observed with a transmission plate of three zones, open aperture center. Both transmission and reflection patterns are shown. Note that the beam width is about 4° in each case. The diameter of this plate was about 16.5 wavelengths. The beamwidth is almost identical with that of a paraboloid having the same diameter. And *this is the salient feature of zone plates*. Even though the gain is much lower than that of a paraboloid of the same diameter (as shown in Fig. 1), the beamwidth is the same. Therefore, in applications where a narrow beamwidth is required and the gain can be achieved by amplification, it becomes feasible to use the zone plate.

The reason for the partial independence of gain and beamwidth is in the very nature of the zone plate as a diffraction device. Instead of scattering all of the incident radiation into the feed like a paraboloid, the zone plate scatters only a portion of it into the feed and hence, has a much lower gain. Beamwidth, on the other hand, is determined by the physical diameter.

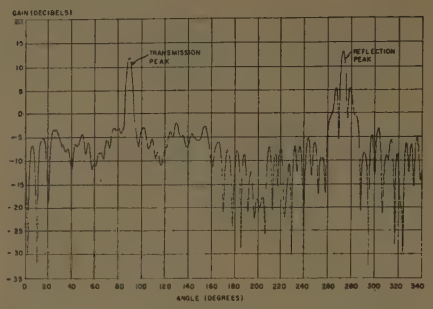



Fig. 3—Zone-plate patterns showing transmission and reflection maxima.

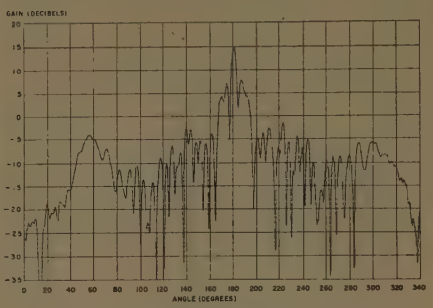



Fig. 4—Pattern of zone plate with ground plane (folded system).

Referring again to Fig. 3, we note that the sidelobe level on transmission is about 15 db, whereas on reflection there is a pair of near sidelobes about 8 db below the peak. These sidelobes have not been explained theoretically, although they may be due to obscuration by the feed antenna.

When a ground plane is added and the transmission and reflection foci are made to coincide, the pattern is as shown in Fig. 4. Here again, near sidelobes appear about 7 db down.

The theoretical gain equations were compared with experimental data by sequentially covering the open zones of a three-zone plate. Table I compares experimental vs theoretical values.

TABLE I
Comparison of Experimental and Theoretical Gains in Relation to Number of Zones (Aperture Center Zone Plate)

| Number of Zones | Theoretical Gain Over Feed (db) | Experimental Gain (db) |
|---|---|---|
| 1 | 6.0 | 4.5 |
| 2 | 11.8 | 9.0 |
| 3 | 15.2 | 11.5 |

### Conclusion

Certain aspects of zone-plate performance have been investigated. Because of the constructional simplicity, it appears that in certain applications where large-aperture focusing elements are required, the zone plate may prove useful. Such applications include low-cost radio telescopes and space communications antennas.

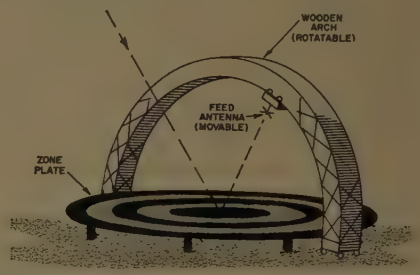



Fig. 5—Fixed zone-plate radio telescope.

Fig. 5 shows a possible configuration based on the folded system. Here the zone plate is fixed in position above the ground plane, and only the feed antenna is moved about. Owing to the deterioration of performance at large scan angles, such a system is limited to rather small angles away from the zenith, perhaps ±20°. However, the cost of constructing such a device would be quite low, compared to even a comparable paraboloid. It is felt that the system might be useful, for example, to a small college desiring to do some work in radio astronomy.

L. F. Van Buskirk
C. E. Hendrix
Michelson Lab.
U. S. Naval Ordnance Test Station
China Lake, Calif

# Contributors

J. T. Bolljahn (A'43–SM'53–F'59), for a photograph and biography, please see this issue, page 234.

❖

Lawrence Brennan (S'47–A'51–M'57) was born in Oak Park, Ill., on January 29, 1927. He received the B.S.E.E. degree in 1948 from the University of Illinois, Urbana, and the Ph.D. degree in 1951 from the same university.

L. E. Brennan

Since October, 1958, he has been a member of the staff of the RAND Corporation, Santa Monica, Calif., where most of his work has been concerned with radar and infrared systems.

Dr. Brennan is a member of Phi Kappa Phi and Sigma Xi.

❖

J. Robert Christian (M'59) was born in New Brunswick, N. J., on July 25, 1933. He received the B.S.E.E. degree from Rutgers University, New Brunswick, N. J., in 1955.

J. R. Christian

Upon graduation he joined the U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N. J., where he was engaged in experimental work on antennas. After serving in the Army in 1957, he rejoined USASRDL and was assigned to the Institute of Exploratory Research. Since then Mr. Christian has been concerned with research and development work on surface wave transmission lines and beam waveguides.

❖

Richard T. Close was born on December 24, 1934, in New York, N. Y. He received the B.S. degree in physics, summa cum laude, from Iona College, New Rochelle, N. Y., in 1956. During the 1956–1957 academic year he was a research fellow at the University of Cincinnati, Ohio. He is now working toward the M.S. degree at the University of Maryland, College Park.

R. T. Close

He served for two years as a Guided Missile Officer in the United States Marine Corps. He was then employed by the Microwave Antennas and Components Branch, Electronics Division, U. S. Naval Research Laboratory, Washington, D. C., where he worked on spiral antennas and their application to large arrays.

Mr. Close is a member of the American Institute of Physics (American Association of Physics Teachers) and the Albertus Magnus Guild.

❖

John R. Donnellan was born on September 2, 1935, in New York, N. Y. He received the B.S. degree in physics from Iona College, New Rochelle, N. Y., in 1956, and then went on to Fordham University, Bronx, N. Y., on an assistantship, where he received the M.S. degree in physics in 1958. He is at present working toward the Ph.D. degree at the University of Maryland, College Park.

J. R. Donnellan

Since 1957, he has been a member of the Microwave Antennas and Components Branch, Electronics Division, U. S. Naval Research Laboratory, Washington, D. C., and has been engaged there in the field of antenna research and development, with particular emphasis on spiral antennas and spiral scanning arrays.

Mr. Donnellan is a member of the American Institute of Physics.

❖

Leonard O. Goldstone (M'51–SM'55) was born in Ogdensburg, N. Y., on February 15, 1922. He received the B.E.E. degree from Clarkson College, Potsdam, N. Y., in 1947, and the M.E.E. and D.E.E. degrees from the Polytechnic Institute of Brooklyn, Brooklyn, N. Y., in 1950 and 1957, respectively.

L. O. Goldstone

From 1947 to 1954, he was an instructor and then assistant professor in the Electrical Engineering Department of the Polytechnic Institute of Brooklyn. From 1954 to 1957, he was a research associate at the Microwave Research Institute at Polytech, where he was engaged in research on propagation in nonconventional waveguides. In 1957 he returned to the faculty of the Institute, where he is presently associate professor of electrical engineering. He is also currently serving as a research consultant for Jasik Laboratories, Inc.

Dr. Goldstone is a member of Eta Kappa Nu and Sigma Xi.

❖

Georg Goubau (A'49–SM'56–F'57) was born in Munich, Germany, on November 29, 1906. He received the M.S. degree in physics in 1930, and the Ph.D degree in physics in 1931, both from the Institute of Technology in Munich.

G. Goubau

From 1931 to 1939, he was engaged in research and teaching in the physics department of the same Institute. In 1939 he was appointed Professor and Director of the department of applied physics of the University of Jena, Germany. He came to this country in 1947, and has been with the USASRDL, Fort Monmouth, N. J., since.

His early work was mainly devoted to ionospheric research. His later research activities include microwave circuit theory, pulse modulation, wave propagation, in particular surface waves, transmission lines and antennas. He is a senior author of the FIAT Review of German Science 1939–1946, and Editor and co-author of a book on electromagnetic waveguides and cavities, published in Germany and recently translated into English by the Atomic Energy Commission.

Dr. Goubau is the recipient of the 1957 Harry Diamond Memorial Award of the IRE and a member of Commission VI of URSI.

❖

Donald E. Johansen (M'57) was born in Los Angeles, Calif., on March 20, 1932. He received the A.B. degree in physics from the University of California, Berkeley, in 1953. Since then he has done graduate work at Massachusetts Institute of Technology, Cambridge, Mass., and is presently a predoctoral student in applied physics at Harvard University.

From 1953 to 1954, he was associated

with Litton Industries (Chromatic Television Laboratories), Emeryville, Calif. On discharge from the Army in 1956, he joined the staff of Hermes Electronics Co., Cambridge, Mass., where he worked in the fields of control and communication theory. Since 1959 he has been employed by Sylvania Electric Products, Inc., Applied Research Laboratory, Waltham, Mass., where he has continued his researches in the general area of communications theory.

D. E. JOHANSEN

❖

Arthur A. Oliner (M'47–SM'52–F'61) was born in Shanghai, China, on March 5, 1921. He received the B.A. degree from Brooklyn College, Brooklyn, N. Y., in 1941. He then attended Cornell University, Ithaca, N. Y., where he received the Ph.D. degree in physics in 1946. While at Cornell, he held a graduate teaching assistantship in the physics department and also conducted research on an Office of Scientific Research and Development project.

A. A. OLINER

Since 1946 he has been with the Microwave Research Institute of the Polytechnic Institute of Brooklyn, where he has conducted research in a wide variety of topics in the microwave field. He has also taught graduate courses in physics and electrical engineering, and is a research professor.

Dr. Oliner is chairman of the IRE Committee on Antennas and Waveguides and is a past chairman of the IRE Professional Group on Microwave Theory and Techniques. He is chairman of USA Commission 1 and a member of Commission 6 of URSI. He also serves as chairman of a National Academy of Sciences Advisory Panel to the National Bureau of Standards.

❖

Felix Schwering (M'60) was born in Cologne, Germany, on June 4, 1930. He attended the University in Aachen, Germany, where he received the M.S. degree in electrical engineering in 1954, and the Ph.D. degree in electrical engineering in 1957.

F. SCHWERING

From 1956 to 1958, he was assistant professor at the Institute of Theoretical Physics at the University in Aachen and performed analytical studies in the fields of electron optics and electromagnetic theory. In 1958 he joined the Institute of Exploratory Research, USASRDL, Fort Monmouth, N. J., where he is currently employed as a physicist in the fields of antenna theory and propagation of electromagnetic waves.

❖

T. L. Simpson was born on December 13, 1935, in Knoxville, Tenn. He received the B.S.E.E. degree in 1956, and the M.S.E.E. degree in 1959, both from University of Tennessee, Knoxville.

T. L. SIMPSON

From 1956 to 1958, he was employed by the Developmental Engineering Corp., Leesburg, Va., where he was engaged in the development of the tapered aperture horn antenna (TAHA). After a year's leave of absence spent in graduate school, Mr. Simpson returned to DECO, where he is currently involved in further research and development of TAHA and other communications antennas.

Mr. Simpson is a member of Tau Beta Pi, Eta Kappa Nu and an associate member of Sigma Xi.

❖

Erling V. Sørensen (A'56) was born in Copenhagen, Denmark, on October 30, 1928. He received the M.S.E.E. degree from the Technical University of Denmark, Copenhagen, in 1953.

E. V. SØRENSEN

From 1955–1957, he worked on microwave and radio equipment with DISA: electronic division of Dansk Industri Syndikat, Comp. Madsen A/S in Copenhagen. During a foreign student summer seminar at M.I.T., Cambridge, Mass., in 1956, he worked on ferrite devices. Since 1957 he has been with the Microwave Laboratory of the Danish Academy of Technical Science in Copenhagen, where he has investigated some properties of microwave ferrites and presently is doing theoretical work on parametric amplifiers.

Mr. Sørensen is a member of the Institution of Danish Civil Engineers.

❖

Seymour Stein (SM'57) was born in Brooklyn, N. Y., on April 4, 1928. He received the B.E.E. degree in 1949 from the College of the City of New York, and the M.S. and Ph.D. degrees in applied physics from Harvard University, Cambridge, Mass., in 1950 and 1955, respectively. In 1951–1952 and 1952–1953, he held an RCA Fellowship in electronics (National Research Council).

From 1953–1956, he was an engineering specialist with Sylvania Electric Products, Inc., Waltham, Mass., primarily engaged in military system studies. From 1956 to 1959, he was with the technical staff at Hermes Electronics Co., Cambridge, working on a variety of problems in antennas, propagation, and communication theory. In 1959 he rejoined Sylvania as a senior engineering specialist, and recently has been appointed senior scientist. He is head of the Communications Research Department of the Applied Research Laboratory, Sylvania Electronic Systems, where he directs and participates in research on a variety of problems with potential communication system applications.

S. STEIN

Dr. Stein is a member of Eta Kappa Nu, Tau Beta Pi, Sigma Xi, and AAAS.

❖

J. D. Tillman, Jr., was born on July 4, 1921, in Evansville, Ind. He received the B.S. degree in 1947, and the M.S. degree in 1950, both in engineering, from the University of Tennessee, Knoxville.

J. D. TILLMAN, JR

He has been on the staff of the University's Electrical Engineering Department since 1947, and is now professor of electrical engineering. Since 1951 he has been engaged mainly in antenna research and development, and since 1956 has been in charge of all antenna work.

Mr. Tillman is a member of Phi Kappa Phi, Eta Kappa Nu, Tau Beta Pi, Sigma Xi, and AIEE.

❖

Jui Lin Yen (S'53–A'54–M'60) was born in Canton, China, on November 29, 1925. He received the B.S. degree in electrical engineering from Chiao Tung University, China, in 1948, and the M.A.S. and Ph.D. degrees in engineering physics from the University of Toronto, Ontario, Canada, in 1950 and 1953, respectively.

J. L. YEN

In 1953 he joined the staff of the University of Toronto, where he is now assistant professor in electrical engineering. He is currently working in antenna theory and radio astronomy.

## PAPERS TO BE PUBLISHED IN FUTURE ISSUES

Reflection of Electromagnetic Waves from a Stratified Inhomogeneity ..... *R. Yamada*

Simultaneous Scintillation Observations on 1300-Mc and 3000-Mc Signals Received during the Solar Eclipse of October 2, 1959 ..... *J. Aarons and J. Castelli*

Frequency Variations Due to Over-the-Horizon Tropospheric Propagation ..... *J. H. Chisholm, S. J. Goodman, J. M. Kennedy, L. B. Lambert, L. P. Rainville, and J. F. Roche*

A Theoretical Limitation on the Formation of Lossless Multiple Beams in Linear Array ..... *J. L. Allen*

Resonance Characteristics of a Corrugated Cylinder Excited by a Magnetic Dipole ..... *J. R. Wait and A. M. Conda*

Studies of Meteor Propagation at 49 and 74 Mc ..... *J. B. Berry, J. C. James, and M. L. Meeks*

The Traveling-Wave Linear Antenna ..... *E. E. Altshuler*

On Propagating Discontinuities in the Electromagnetic Field ..... *K. R. Johnson*

A Method for Computing Ionospheric Focusing of Radio Waves, Using Vertical Incidence Ionograms ..... *E. Warren and D. Muldrew*

New Circularly-Polarized Frequency Independent Antennas with Conical Beam or Omnidirectional Patterns ..... *J. D. Dyson and P. E. Mayes*

A General Analysis of Nonplanar, Two-Dimensional Luneberg Lenses ..... *S. Adachi, R. C. Rudduck, and C. H. Walter*

The Numerical Evaluation of Radiation Integrals ..... *J. H. Richmond*

Arbitrarily-Polarized Planar Antennas ..... *F. J. Goebels, Jr. and K. C. Kelly*

Elevated Duct Propagation in the Tradewinds ..... *F. C. Macdonald*

An Iris-Excited Slot Radiator in the Narrow Wall of Rectangular Waveguide ..... *D. G. Dudley*

Dipole Antennas Coupled Electromagnetically to a Two-Wire Transmission Line ..... *K. M. Chen and R. W. P. King*

A New Way of Solving Maxwell's Equations ..... *V. H. Rumsey*

Correlation of Wind Shear with Tropospheric Scatter Signals ..... *L. H. Bauer*

# INSTITUTIONAL LISTINGS

The IRE Professional Group on Antennas and Propagation is grateful for the assistance given by the firms listed below, and invites application for Institutional Listing from other firms interested in the field of Antennas and Propagation.

AERO GEO ASTRO CORP., 1200 Duke St., Alexandria, Va.
Space Instrumentation; Antennas; Transponders; Command Receivers; Augmenters; Telemetry; Radar

ANDREW CORPORATION, P.O. Box 807, Chicago 42, Ill.
Antennas, Antenna Systems, Transmission Lines, Development and Production

ANTLAB, INC., 6330 Proprietors Rd., Worthington, Ohio
Antenna Pattern Range Systems—Recorders & Mounts, & Telemetry Servo Pedestals

BLAINE ELECTRONETICS, INC., 14757 Keswick St., Van Nuys, Calif.
Antennas, Paraboloids, Scale Models, Antenna Radiation Pattern Measurement Towers

DEVELOPMENTAL ENGINEERING CORP., Leesburg, Va.; Boston, Mass.; Boulder, Colo.; Washington, D.C
Antenna Research, Design & Evaluation-Propagation Studies & Communications Systems Engineering

DORNE & MARGOLIN, INC., 29 New York Ave., Westbury, L. I., N. Y.
Research, Development and Manufacture—Antenna and Microwave Technology

FXR, Inc., 25–26 50th St., Woodside 77, N. Y.
Precision Microwave Test Equip., High Power Microwave Electronics, Microwave Components & Instrumentatio

HUGHES AIRCRAFT COMPANY, Florence and Teale Sts., Culver City, Calif.
Res., Dev., Mfg.: Radar Systems & Components; Antennas, Tubes, Radomes, Solid-State Devices

MARK PRODUCTS CO., 5439 W. Fargo Ave., Skokie, Ill.
Antennas for Two-Way Communications, Grid Parabolas, Research & Development

SCIENTIFIC-ATLANTA, INC., 2162 Piedmont Rd., N.E., Atlanta 9, Ga.
Antenna Pattern Ranges, R & D, RF Components, Telemetry Tracking Systems, Fourier Computer & Computations

TECHNICAL APPLIANCE CORP., 1 Taco St., Sherburne, N. Y.
Des., Dev., & Mfg.: Antennas & Antenna Systems for Communications, Telemetry, & Tracking

WEINSCHEL ENGINEERING COMPANY, INC., Kensington, Md.
Antenna Pattern Receivers; Bolometer Amplifiers; Microwave Sources; Insertion Loss Measuring Systems

WHEELER LABORATORIES, INC., Great Neck, N. Y.; Antenna Lab., Smithtown, N. Y.
Consulting Services, Research and Development, Microwave Antennas and Waveguide Components

The charge for Institutional Listing is $50 for one issue or $200 for six consecutive issues (one year). Application may be made to the Technical Secretary, The Institute of Radio Engineers, Inc., 1 East 79th Street, New York 21, N. Y.